PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : G01N 33/574, C07K 16/30, C12Q 1/68, G01N 33/561 — A1

(11) International Publication Number: WO 00/55628

(43) International Publication Date: 21 September 2000 (21.09.00)

(21) International Application Number: PCT/GB00/00908

(22) International Filing Date: 13 March 2000 (13.03.00)

(30) Priority Data:
9905817.4 12 March 1999 (12.03.99) GB

(71) Applicant *(for all designated States except US)*: OXFORD GLYCOSCIENCES (UK) LTD. [GB/GB]; 10 The Quadrant, Abingdon Science Park, Abingdon, Oxfordshire OX14 3YS (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: AMESS, Bob [GB/GB]; Oxford GlycoSciences Ltd., 10 The Quadrant, Abingdon Science Park, Abingdon, Oxfordshire OX14 3YS (GB). TOWNSEND, Robert, Reid [GB/GB]; Oxford GlycoSciences Ltd., 10 The Quadrant, Abingdon Science Park, Abingdon, Oxfordshire OX14 3YS (GB). PAREKH, Rajesh Bhikhu [GB/GB]; Oxford GlycoSciences Ltd., 10 The Quadrant, Abingdon Science Park, Abingdon, Oxfordshire OX14 3YS (GB). WATERFIELD, Michael, Derek [GB/GB]; Ludwig Institute for Cancer Research, 91 Riding House Street, London W1P 8BT (GB). O'HARE, Michael, John [GB/GB]; Ludwig Institute For Cancer Research, 91 Riding House Street, London W1P 8BT (GB).

(74) Agents: CHAPMAN, Paul, William et al.; Kilburn & Strode, 20 Red Lion Street, London WC1R 4PJ (GB).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PROTEINS FOR DIAGNOSIS AND TREATMENT OF BREAST CANCER

(57) Abstract

The present invention provides methods and compositions for screening, diagnosis and prognosis of breast cancer, for monitoring the effectiveness of breast cancer treatment, and for drug development. Breast Cancer–Associated Features (BFs), detectable by two–dimensional electrophoresis of breast tissue are described. The invention further provides Breast Cancer–Associated Protein Isoforms (BPIs) detectable in breast tissue, preparations comprising isolated BPIs, antibodies immunospecific for BPIs, and kits comprising the aforesaid.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

PROTEINS FOR DIAGNOSIS AND TREATMENT OF BREAST CANCER

1. INTRODUCTION

The present invention relates to the identification of proteins and protein isoforms that are associated with the onset and progression of breast cancer and to their use for screening, diagnosis, prognosis, therapy and drug development.

2. BACKGROUND OF THE INVENTION

Despite major advances in basic research involving molecular/cell biology and signal transduction over the last two decades, breast cancer continues to be a leading cause of death for women. Landis et al. (1998) CA Cancer J Clin 48: 6-29; Cancer Research Campaign (UK) (1996) Factsheet 6.1. Response rates to therapy remain poor, and the period of survival for women who present with metastatic disease is typically only between 18-24 months. Leonard et al.. (1994) BMJ 309: 1501-1504. Although many prognostic indicators have been forwarded, there nevertheless remains a significant degree of uncertainty for breast cancer patients in the course their disease will take. Ravaioli et al.(1998) Cell Prolif. 31:113-126; Ferno, M. (1998) Anticancer Res. 18:2167-2171. This may be attributable to over-reliance on a limited number of genetic and proliferative indicators, which is compounded by the extensive degree of heterogeneity present in most advanced tumours.

A preferred situation would be the ability to identify simultaneously all of the changes in a tumour, which occurred at the gene or protein level. This is being addressed by two new areas of biology termed genomics and proteomics. The area of genomics has advanced such that gene mutations and transcript mRNAs can now be characterised in a high throughput manner. Crooke (1998) Nat. Biotechnol. 16:29-30; Schena et al. (1998) Trends Biotechnol. 16:301-306; Ramsay (1998) Nat. Biotechnol. 16:40-44. To complement this, proteomics, which examines the entire protein repertoire of a sample, and can identify post-translational modifications and sub-cellular localisation, has now reached a similar level of refinement. Anderson & Anderson (1998) Electrophoresis

19:1853-1861; James (1997) Biochem. Biophys. Res. Commun. 231:1-6; Page et al. (1999) Drug Discovery Today 4:55-62. Each technology is capable of identifying in a single process, tens or hundreds of events at the gene or protein level, which have changed in a tumour. These tumour-associated changes will represent new classes of markers, or even new candidate targets for therapeutic intervention. By clustering sets of these new markers, with drug response, and ultimately survival, it is hoped that clinicians will have powerful new diagnostic and prognostic information that will be of benefit to their patients.

A critical part of this process as its applies to breast and other cancers, is the ability to identify and define the changes which have occurred between normal and tumour material. The human breast is a complex organ whose proliferative and differentiative capacity is tightly regulated by the interplay of growth factors, steroid hormones and cell-cell interactions. Dickson et al. (1993) Adv. Exp. Med. Biol. 330:119-141; Rudland et al. (1998) Biochem. Soc. Symp. 63:1-20. The terminal lobular-alveolar unit of the breast, which is the structure from which the majority of cancers arise, is composed of two types of epithelial cells. The inner or luminal epithelial cells are potential milk secretory cells which are surrounded by an outer layer of contractile myoepithelial cells. Most of breast carcinomas (95%) express phenotypic markers, which are consistent with an origin from luminal rather than myo-epithelial cells. Taylor-Papadimitriou and Lane (eds. M.C. Neville and C.W. Daniel) pp.181-215 Plenum Press: New York (1987).

Attempts to resolve proteins by 2-dimensional polyacylamide gel electrophoresis (2-D PAGE) from breast material begun in the mid-late 1980's and were considerably hampered by the lack of reproducibility, poor sensitivity, use of biopsy material containing mixed cell populations, and lack of identification of the proteins of interest. Stastny et al. (1984) Clin. Chem. 30:1914-1918; Wirth et al. (1987) Breast Cancer Res. Treat. 10:177-189; Wirth (1989) Electrophoresis 10:543-554. More recently due to the significant progress made in key areas of proteomics, there are now reports and World Wide Web databases describing proteomes for normal and tumour derived breast cell lines in culture and breast tissue from benign and malignant sources. Figeys et al. (1998) Electrophoresis 19:1811-1818; Corthals et al. (1997) Electrophoresis 18:317-323; Giometti et al. (1997) Electrophoresis 18:573-581; Williams et al. (1998) Electrophoresis

19:333-343; Rasmussen et al. (1998) Electrophoresis 19:818-825; Bini et al. (1997) Electrophoresis 18:2832-2841; Franzen et al. (1997) Electrophoresis 18:582-587. However, these studies are still limited by a number of factors: the extensive use of established cell lines which may be unrepresentative of either normal luminal or malignant breast, the use of normal cells of undefined phenotype and the heterogeneity of cell types where primary breast material has been studied. Notably there are no published proteomes for purified normal human breast cells.

It is now possible, using immunomagnetic methods, to purify the luminal and myoepithelial cells from normal human breast material, in sufficient quantities to enable detailed analysis, including proteomic characterisation. Clarke et al. (1994) Epithelial Cell Biol. 3:38-46; Gomm et al. (1995) Analytical Biochem. 226:91-99. We have now completed the first complete proteome analysis of ten sets of matched purified normal human luminal and myoepithelial breast cell populations obtained directly from reduction mammoplasties and purified by a double selection method that results in purities equivalent to those obtained by FACS sorting. O'Hare et al.. (1991) Differentiation 46:209-221. Furthermore, high resolution detection of protein features using fluorescent dyes, coupled to advanced software to identify differentially expressed features, high through-put mass spectrometry and bioinformatics has also been applied. This has allowed the identification of large sets of proteins which are differentially expressed between the luminal and myoepithelial human breast cell proteomes.

Both the constant and differential sets of protein markers identified herein can therefore be used to detect the presence of breast cells. Thus, as described below, we have now provided a means for the detection and /or diagnosis of metastatic breast tumour cells, based on detection of one or more of the markers, or more usually, based on identification of a cluster of such markers.

3. SUMMARY OF THE INVENTION

The present invention provides methods and compositions for screening, diagnosis, prognosis and therapy of breast cancer, for monitoring the effectiveness of breast cancer treatment, and for drug development.

Thus, in various aspects, the present invention provides,

a method for screening for and/or diagnosis of breast cancer in a human subject, which comprises the step of identifying the presence or absence of one or more of the protein features as defined in tables I, II, III and IV herein, in a biological sample obtained from said human subject;

a method for monitoring and/or assessing breast cancer treatment in a human subject, which comprises the step of identifying the presence or absence of one or more of the protein features as defined in tables I, II, III and IV herein, in a biological sample obtained from said human subject; and

a method for identifying the presence or absence of metastatic breast cancer cells in a biological sample obtained from a human subject, which comprises the step of identifying the presence or absence of one or more of the protein features as defined in tables I, II, III and IV herein.

In additional aspects, the present invention provides:

methods for diagnosis of breast cancer that comprise analyzing a sample of tissue by two-dimensional electrophoresis to detect the level of at least one breast cancer-associated feature (BF), e.g., one or more of the BFs disclosed herein or any combination thereof. Especially preferred is a method that comprises quantitative detection of a cluster of BFs, such as the clusters of BFs described herein (these methods are also suitable for screening, prognosis, monitoring the results of therapy, and drug development);

methods for diagnosis of breast cancer that comprise detecting in a sample of tissue the level of at least one breast cancer-associated protein isoform (BPI), e.g., one or more of the BPIs disclosed herein or any combination thereof. Especially preferred is a method that comprises quantitative detection of a cluster of BPIs, such as the clusters of BPIs described herein (these methods are also suitable for screening, prognosis, monitoring the results of therapy, and drug development);

monoclonal and polyclonal antibodies capable of immunospecific binding to a BPI, e.g., a BPI disclosed herein;

a preparation comprising an isolated BPI, i.e., a BPI free from proteins or protein isoforms having a significantly different isoelectric point or a significantly different apparent molecular weight from the BPI; and

an antibody or antibodies capable of immunospecific binding to a BPI, e.g., a BPI disclosed herein, for systemic therapy or other therapies localised to sites of breast cancer metastasis.

4. DETAILED DESCRIPTION OF THE INVENTION

The invention described in detail below encompasses methods and compositions for screening, diagnosis and prognosis of metastatic breast cancer in a subject, for monitoring the results of breast cancer therapy, and for drug development. The invention also encompasses the administration of therapeutic compositions to a mammalian subject to treat or prevent breast cancer. Preferably, the mammalian subject is human, more preferably a human adult. For clarity of disclosure, and not by way of limitation, the invention will be described with respect to the analysis of breast tissue samples. However, as one skilled in the art will appreciate, the assays and techniques described below can be applied to other types of patient samples, including a body fluid (e.g. blood, serum, plasma or saliva), a tissue sample from a patient at risk of having metastatic breast cancer (e.g. a biopsy such as a breast biopsy) or homogenate thereof. The methods and compositions of the present invention are specially suited for screening, diagnosis and prognosis of a living subject, but may also be used for postmortem diagnosis in a subject, for example, to identify family members at risk of developing the same disease.
As used herein, breast tissue refers to the tissue within the breast and may also include tissue taken from the strata underlying or adjacent to the breast.

4.1. Breast Cancer-Associated Features (BFs)

In one aspect of the invention, two-dimensional electrophoresis is used to analyze breast tissue from a subject, preferably a living subject, in order to measure the expression of one or more Breast Cancer-Associated Features (BFs) for screening or diagnosis of breast

cancer, to determine the prognosis of an breast cancer patient, to monitor the effectiveness of breast cancer therapy, or for drug development. As used herein, "two-dimensional electrophoresis" (2D-electrophoresis) means a technique comprising isoelectric focusing, followed by denaturing electrophoresis; this generates a two-dimensional gel (2D-gel) containing a plurality of separated proteins. Preferably, the step of denaturing electrophoresis uses polyacrylamide electrophoresis in the presence of sodium dodecyl sulfate (SDS-PAGE). Especially preferred are the highly accurate and automatable methods and apparatus ("the Preferred Technology") described in U.S. Application No. 08/980,574 filed December 1, 1997 (published as WO 98/23950), which is incorporated herein by reference in its entirety with particular reference to the preferred protocol described therein. Briefly, the Preferred Technology provides efficient, computer-assisted methods and apparatus for identifying, selecting and characterizing biomolecules in a biological sample. A two-dimensional array is generated by separating biomolecules on a two-dimensional gel according to their electrophoretic mobility and isoelectric point. A computer-generated digital profile of the array is generated, representing the identity, apparent molecular weight, isoelectric point, and relative abundance of a plurality of biomolecules detected in the two-dimensional array, thereby permitting computer-mediated comparison of profiles from multiple biological samples, as well as computer aided excision of separated proteins of interest.

A preferred scanner for detecting fluorescently labeled proteins is described in WO 96/36882 and in the Ph.D. thesis of David A. Basiji, entitled "Development of a High-throughput Fluorescence Scanner Employing Internal Reflection Optics and Phase-sensitive Detection (Total Internal Reflection, Electrophoresis)", University of Washington (1997), Volume 58/12-B of Dissertation Abstracts International, page 6686, the contents of each of which are incorporated herein by reference. These documents describe an image scanner designed specifically for automated, integrated operation at high speeds. The scanner can image gels that have been stained with fluorescent dyes or silver stains, as well as storage phosphor screens. The scanner provides a phase-sensitive detection system for discriminating modulated fluorescence from baseline noise due to laser scatter or homogeneous fluorescence, but can also be operated in a non-phase-sensitive mode. This phase-sensitive detection capability increases the sensitivity of the instrument by an order of magnitude or more compared to conventional fluorescence imaging systems. The increased sensitivity reduces the sample-preparation load on the

upstream instruments while the enhanced image quality simplifies image analysis downstream in the process.

A more highly preferred scanner is the Apollo 2 scanner (Oxford GlycoSciences, Oxford, UK), which is a modified version of the above described scanner. In the Apollo 2 scanner, the gel is transported through the scanner on a precision lead-screw drive system. This is preferable to laying the glass plate on the belt-driven system that is defined in the Basiji thesis as it provides a reproducible means of accurately transporting the gel past the imaging optics.

In the Apollo 2 scanner, the gel is secured against three alignment stops that rigidly hold the glass plate in a known position. By doing this in conjunction with the above precision transport system, the absolute position of the gel can be predicted and recorded. This ensures that co-ordinates of each feature on the gel can be determined more accurately and communicated, if desired, to a cutting robot for excision of the feature. In the Apollo 2 scanner, the carrier that holds the gel has four integral fluorescent markers used to correct the image geometry. These markers are a quality control feature to confirm that the scanning has been performed correctly.

In comparison to the scanner described in the Basiji thesis, the optical components of the Apollo 2 scanner have been inverted. In the Apollo 2 scanner, the laser, mirror, waveguide and other optical components are above the glass plate being scanned. The scanner described in the Basiji thesis has these components underneath. In the Apollo 2 scanner, the glass plate is mounted onto the scanner gel side down, so that the optical path remains through the glass plate. By doing this, any particles of gel that may break away from the glass plate will fall onto the base of the instrument rather than into the optics. This does not affect the functionality of the system, but increases its reliability.

As used herein, the term "Breast Cancer-Associated Feature" (BF) refers to a feature (e.g., a spot in a 2D gel), detectable by 2D electrophoresis of a breast tissue sample, that is differentially present in sets of matched purified normal human luminal cells, compared with corresponding purified normal human myoepithelial cells. As used herein, a feature (or a protein isoform) is "differentially present" in a first sample with respect to a second sample when a method for detecting the feature or isoform (e.g., 2D electrophoresis or an immunoassay) gives a different signal when applied to the first and second samples. If the first sample gives a higher signal than the second sample, the feature or isoform is "increased" in the first sample with respect to the second;

conversely, if the first sample gives a lower signal than the second sample, the feature or isoform is "decreased" in the first sample with respect to the second.

Preferably, the relative abundance of a feature in two samples is determined in two steps. First, the signal obtained upon detecting the feature in a sample is normalized by reference to a suitable background parameter, e.g., (a) to the total protein in the sample being analyzed (e.g., total protein loaded onto a gel); (b) to an invariant feature, i.e., a feature whose abundance is known to be similar in the samples being compared; or (c) to the total signal detected from all proteins in the sample.

Secondly, the normalized signal for the feature in one sample or sample set is compared with the normalized signal for the same feature in another sample or sample set in order to identify features that are "differentially present" in the first sample (or sample set) with respect to the second.

The BFs disclosed herein have been identified by comparing sets of matched normal breast luminal cells, purified from normal breast tissue of living human subjects, with corresponding sets of matched purified normal human myoepithelial cells from the same subjects. The BFs are those features that are differentially present in sets of matched normal breast luminal cells, purified from normal breast tissue of living human subjects, as compared with corresponding sets of matched purified normal human myoepithelial cells from the same subjects. Subjects free from breast cancer include subjects with no known disease or condition (normal subjects) and subjects with diseases other than breast cancer.

Two groups of BFs have been identified through the methods and apparatus of the Preferred Technology. The first group consists of BFs that are decreased in sets of matched normal breast luminal cells, purified from normal breast tissue of living human subjects, as compared with corresponding sets of matched purified normal human myoepithelial cells from the same subjects. These BFs can be described by apparent molecular weight (MW) and isoelectric point (pI) as follows:

Table I. BFs decreased in normal breast luminal epithelial cells

BF	Fold Decrease	pI	Mr (kDa)
BF-1	44.4	5.48	27415

BF-2	25.0	4.56	19809
BF-3	19.0	5.01	30039
BF-4	17.3	6.57	56041
BF-5	17.3	6.63	57276
BF-6	15.6	6.79	56216
BF-7	15.0	4.75	27415
BF-8	13.8	4.61	25841
BF-9	13.7	5.91	24964
BF-10	13.6	4.76	16387
BF-11	13.4	6.15	55175
BF-12	13.3	7.84	54833
BF-13	12.9	4.71	20575
BF-14	11.1	4.79	28644
BF-15	10.7	7.91	41347
BF-16	10.1	9.27	15777
BF-17	9.6	5.11	27897
BF-18	9.2	4.70	27550
BF-19	9.1	4.79	21533
BF-20	7.4	5.64	26096
BF-21	7.3	4.66	23406
BF-22	6.5	7.04	39902
BF-23	6.3	7.32	48045
BF-24	5.6	7.33	33743
BF-25	4.2	8.63	28644
BF-26	3.6	5.46	54833
BF-27	3.1	5.41	52330
BF-28	2.7	5.43	55175
BF-29	2.4	5.43	36557
BF-30	2.3	4.90	51362
BF-31	2.0	9.79	33921
BF-32	2.0	5.27	31171

BF-33	2.0	6.03	43639

The second group consists of BFs that are increased in sets of matched normal breast luminal cells, purified from normal breast tissue of living human subjects, as compared with corresponding sets of matched purified normal human myoepithelial cells from the same subjects. These BFs can be described by apparent molecular weight (MW) and isoelectric point (pI) as follows:

Table II. BFs increased in normal breast luminal epithelial cells

BF	Fold Increase	pI	Mr (kDa)
BF-34	28.7	9.14	37133
BF-35	23.7	4.66	32525
BF-36	22..0	5.30	35165
BF-37	19.6	5.57	14596
BF-38	16.8	7.59	13120
BF-39	15.4	8.11	58407
BF-40	14.7	6.48	21250
BF-41	13.3	6.99	31614
BF-42	13.3	6.43	23412
BF-43	10.4	8.66	48341
BF-44	8.1	4.99	70465
BF-45	3.9	4.63	41917
BF-46	3.4	8.61	57516
BF-47	3.1	5.40	71346
BF-48	2.8	7.85	37264
BF-49	2.8	5.28	36758
BF-50	2.3	5.44	34752
BF-51	2.0	5.27	31764

For any given BF, the signal obtained upon analyzing metastatic cancer tissue or cells from subjects having breast cancer will depend upon the particular analytical protocol and detection technique that is used. Accordingly, the present invention contemplates that each laboratory will establish a reference range for each BF in subjects free from breast cancer according to the analytical protocol and detection technique in use, as is conventional in the diagnostic art. Preferably, at least one control positive metastatic cancer tissue sample from a subject known to have breast cancer or at least one control negative tissue sample from the same tissue site in a subject known to be free from breast cancer (and more preferably both positive and negative control samples) are included in each batch of test samples analyzed. In one embodiment, the level of expression of a feature is determined relative to a background value, which is defined as the level of signal obtained from a proximal region of the 2D gel image that (a) is equivalent in area to the particular feature in question; and (b) contains no discernable protein feature.
As those of skill in the art will readily appreciate, the measured MW and pI of a given feature or protein isoform will vary to some extent depending on the precise protocol used for each step of the 2D electrophoresis and for landmark matching. As used herein, the terms "MW" and "pI" are defined, respectively, to mean the apparent molecular weight and the isoelectric point of a feature or protein isoform as measured in exact accordance with the experimental protocol set forth in Section 5 below ("the Reference Protocol"). When the Reference Protocol is followed and when samples are run in duplicate or a higher number of replicates, variation in the measured mean pI of a BF is typically less than 1% and variation in the measured mean MW of a BF is typically less than 5%. Where the skilled artisan wishes to deviate from the Reference Protocol, calibration experiments should be performed to compare the MW and pI for each BF or protein isoform as detected (a) by the Reference Protocol and (b) by the deviant protocol.
BFs can be used for detection, prognosis, diagnosis, or monitoring of breast cancer or for drug development. In one embodiment of the invention, metastatic cancer tissue or cells from a subject (e.g., a subject suspected of having breast cancer) is analyzed by 2D electrophoresis for quantitative detection of one or more of the following BFs: BF-1 to BF-33. A decreased abundance of said one or more BFs in metastatic cancer tissue or cells from the subject relative to tissue or cells from the same tissue site in a subject or subjects free from breast cancer (e.g., a control sample or a previously determined reference range) indicates the presence of breast cancer.

In another embodiment of the invention, metastatic cancer tissue or cells from a subject (e.g., a subject suspected of having breast cancer) is analyzed by 2D electrophoresis for quantitative detection of one or more of the following BFs: BF-34 to BF-51. An increased abundance of said one or more BFs in metastatic cancer tissue or cells from the subject relative to tissue or cells from the same tissue site in a subject or subjects free from breast cancer (e.g., a control sample or a previously determined reference range) indicates the presence of metastatic breast cancer.

In a preferred embodiment, tissue from a subject is analyzed for quantitative detection of clusters of BFs.

4.2. Breast Cancer-Associated Protein Isoforms (BPIs)

In another aspect of the invention, breast tissue from a subject, preferably a living subject, is analyzed for quantitative detection of one or more Breast Cancer-Associated Protein Isoforms (BPIs) for screening or diagnosis of breast cancer, to determine the prognosis of a breast cancer patient, to monitor the effectiveness of breast cancer therapy, or for drug development. As used herein, the term "Breast Cancer-Associated Protein Isoform" refers to a protein isoform that is differentially present in metastatic cancer tissue or cells from a subject having metastatic breast cancer compared with tissue from the same site in a subject free from breast cancer. As is well known in the art, a given protein may be expressed as variants (isoforms) that differ in their amino acid composition (e.g., as a result of alternative splicing or limited proteolysis) or as a result of differential post-translational modification (e.g., glycosylation, phosphorylation, acylation), or both, so that proteins of identical amino acid sequence can differ in their pI, MW, or both. It follows that differential presence of a protein isoform does not require differential expression of the gene encoding the protein in question.

Two groups of BPIs have been identified by partial amino acid sequencing of BFs, using the methods and apparatus of the Preferred Technology. The first group consists of BPIs that are decreased in sets of matched normal breast luminal cells, purified from normal breast tissue of living human subjects, as compared with corresponding sets of matched purified normal human myoepithelial cells from the same subjects. The MW, pI and partial amino acid sequences of these BPIs are presented in Table III, as follows:

Table III. BPIs decreased in normal breast luminal epithelial cells

BF#	BPI	Known homologous protein	Partial Amino Acid Sequence	pI	Mr (kDa)
1	BPI-1	P08727 keratin, type i cytoskeletal 19	(R)DYSHYYTTIQDLR (R)IVLQIDNAR (R)QSSATSSFGGLGG GSVR (K)NHEEEISTLR	5.48	27415
2	BPI-2	P05783 keratin, type i cytoskeletal 18	LEAEIATYR	4.56	19809
3	BPI-3	P08779 keratin, type i cytoskeletal 17	(R)VLDELTLAR (R)TKFETEQALR (R)LSGGLGAGSCR	5.01	30039
4	BPI-4	P02538 keratin, type ii cytoskeletal 6a	(R)QLDSIVGER (K)QEIAEINR (R)ISIGGGSCAISGGY GSR (R)AEAESWYQTK (R)TAAENEFVTLK (K)QCANLQAAIADAE QR (K)LALDVEIATYR (R)SGFSSVSVSR	6.57	56041
5	BPI-5	P07355 annexin ii (lipocortin ii)	QDIAFAYQR	6.63	57276
6	BPI-6	P13647 keratin, type ii cytoskeletal 5	(K)YEELQQTAGR (K)AQYEEIANR (R)LRAEIDNVKK (R)VSLAGACGVGGY GSR (R)SFSTASAITPSVSR ISISTSGGSFR	6.79	56216
7	BPI-7	P08779 keratin, type i cytoskeletal 17	(R)VLDELTLAR (R)TKFETEQALR (R)LSGGLGAGSCR	4.75	27415
8	BPI-8	P08779 keratin, type i cytoskeletal 17	(R)VLDELTLAR (R)TKFETEQALR (R)LSGGLGAGSCR	4.61	25841
9	BPI-9	P08727 keratin, type i cytoskeletal 19	(R)DYSHYYTTIQDLR (R)IVLQIDNAR (R)QSSATSSFGGLGG GSVR (K)NHEEEISTLR	5.91	24964
10	BPI-10	P08727 keratin, type i cytoskeletal 19	(R)DYSHYYTTIQDLR (R)IVLQIDNAR (R)QSSATSSFGGLGG GSVR (K)NHEEEISTLR	4.76	16387
11	BPI-11	P02538 keratin, type ii cytoskeletal 6a	(R)QLDSIVGER (K)QEIAEINR (R)ISIGGGSCAISGGY GSR (R)AEAESWYQTK (R)TAAENEFVTLK	6.15	55175

			(K)QCANLQAAIADAE QR (K)LALDVEIATYR (R)SGFSSVSVSR		
12	BPI-12	P02538 keratin, type ii cytoskeletal 6a	(R)QLDSIVGER (K)QEIAEINR (R)ISIGGGSCAISGGY GSR (R)AEAESWYQTK (R)TAAENEFVTLK (K)QCANLQAAIADAE QR (K)LALDVEIATYR (R)SGFSSVSVSR	7.84	54833
13	BPI-13	P08779 keratin, type i cytoskeletal 17	(R)VLDELTLAR (R)TKFETEQALR (R)LSGGLGAGSCR	4.71	20575
14	BPI-14	P08779 keratin, type i cytoskeletal 17	(R)VLDELTLAR (R)TKFETEQALR (R)LSGGLGAGSCR	4.79	28644
15	BPI-15	P12532 creatine kinase (ec 2.7.3.2)	(K)SFLIWVNEEDHTR (R)LYPPSAEYPDLR	7.91	41347
16	BPI-16	O15509 p20-arc (arp2/3 complex)	(K)ELLLQPVTISR (K)VLIEGSINSVR (R)IVAEEFLK	9.27	15777
17	BPI-17	P08727 keratin, type i cytoskeletal 19	(R)DYSHYYTTIQDLR (R)IVLQIDNAR (R)QSSATSSFGGLGG GSVR (K)NHEEEISTLR	5.11	27897
18	BPI-18	P08727 keratin, type i cytoskeletal 19	(R)DYSHYYTTIQDLR (R)IVLQIDNAR (R)QSSATSSFGGLGG GSVR (K)NHEEEISTLR	4.70	27550
19	BPI-19	P08779 keratin, type i cytoskeletal 17	(R)VLDELTLAR (R)TKFETEQALR (R)LSGGLGAGSCR	4.79	21533
20	BPI-20	P08727 keratin, type i cytoskeletal 19	(R)DYSHYYTTIQDLR (R)IVLQIDNAR (R)QSSATSSFGGLGG GSVR (K)NHEEEISTLR	5.64	26096
21	BPI-21	P08779 keratin, type i cytoskeletal 17	(R)VLDELTLAR (R)TKFETEQALR (R)LSGGLGAGSCR	4.66	23406
22	BPI-22	P04075 fructose-bisphosphate aldolase a	(R)LSQIGVENTEENR (K)ELSDIALR (K)DDNGVPFVR	7.04	39902
23	BPI-23	P13645 keratin, type i cytoskeletal 10	(R)LKYENEVALR (R)SLLEGEGSSGGGG R (K)GSLGGGFSSGGFS	7.32	48045

			GGSFSR (K)DAEAWFNEK (R)QSVEADINGLR (K)YENEVAL		
24	BPI-24	P48739 phosphatidylinositol transfer protein	(K)TVEIVHIDIAD	7.33	33743
25	BPI-25	P13645 keratin, type i cytoskeletal 10	(R)LKYENEVALR (R)SLLEGEGSSGGGG R (K)GSLGGGFSSGGFS GGSFSR (K)DAEAWFNEK (R)QSVEADINGLR (K)YENEVAL	8.63	28644
26	BPI-26	P05787 keratin, type ii cytoskeletal 8	(K)LLEGEESR TEISEMNR	5.46	54833
27	BPI-27	P05787 keratin, type ii cytoskeletal 8	(K)LLEGEESR TEISEMNR	5.41	52330
28	BPI-28	P08729 keratin, type ii cytoskeletal 7	(R)EVTINQSLLAPLR (K)QEELEAALQR (K)YEDEINR EYQELMNVKDVDEA YMNK	5.43	55175
29	BPI-29	P52907 f-actin capping protein alpha-1	(R)LLLNNDNLLR	5.43	36557
30	BPI-30	P02533 keratin, type i cytoskeletal 14	(K)EVASNSELVQSSR (K)ASLEGNLAETENR (R)ISSVLAGGSCR (K)NHEEEMNALR (K)DAEEWFFTK	4.90	51362
31	BPI-31	P22626 hnrnp a2 and hnrnp b1	(K)IDTIEIITDR (K)LFIGGLSFETTEES LR (R)GGGGNFGPGPGSN FR (K)YHTINGHNAEVR (R)DYFEEYGK	9.79	33921
32	BPI-32	P07339 cathepsin d precursor (ec 3.4.23.5).	(R)YYTVFDR (K)LSPEDYTLK	5.27	31171
33	BPI-33	Q92524 26s protease regulatory subunit 10b	(K)IHIDLPNEQAR (R)EVIELPLTNPELFQ R (K)LSDGFNGADLR (K)HGEIDYEAIVK	6.03	43639

The second group comprises BPIs that are increased in sets of matched normal breast luminal cells, purified from normal breast tissue of living human subjects, as compared with corresponding sets of matched purified normal human myoepithelial cells from the

same subjects. The MWs, pIs and partial amino acid sequences of these BPIs are presented in Table IV, as follows:

Table IV. BPIs increased in normal breast luminal epithelial cells

BF#	BPI	Known homologous protein	Partial Amino Acid Sequence	pI	Mr (kDa)
34	BPI-34	P09972 fructose-bisphosphate aldolase c	(R)LSQIGVENTEENR (K)ELSDIALR (K)DDNGVPFVR	9.14	37133
35	BPI-35	P07226 tropomyosin, fibroblast type (tm30-pl)	(K)HIAEEADR (R)IQLVEEELDR IQVLQQQADDAEER	4.66	32525
36	BPI-36	P02570 beta-actin, cytoplasmic 1	(R)GYSFTTTAER	5.30	35165
37	BPI-37	P37802 sm22-alpha homolog (kiaa0120).	(R)NFSDNQLQEGK (K)DVGRPQPGR	5.57	14596
38	BPI-38	P09455 retinol-binding protein i, cellular	(K)YDEELEER (K)EFEEDLTGIDDR	7.59	13120
39	BPI-39	P14618 pyruvate kinase, muscle specific isozyme (ec 2.7.1.40)	(K)GDYPLEAVR (R)LDIDSPPITAR	8.11	58407
40	BPI-40	Q06830 thioredoxin peroxidase 2	QITVNDLPVGR	6.48	21250
41	BPI-41	O15144 arp2/3 complex 34 kd subunit	(R)DNTINLIHTFR (K)ELQAHGADELLK (K)YFQFQEEGK	6.99	31614
42	BPI-42	P04792 heat shock 27 kd protein (hsp 27)	LLKDFFNGK CNEIINWLDK STAGDTHLGGEDFDNR (K)DAGTIAGLNVLR (R)IINEPTAAAIAYGLD K (R)TTPSYVAFTDTER (R)VEIIANDQGNR (K)EEFEHQQK (K)SQIHDIVLVGGSTR (R)FEELNADLFR	6.43	23412
43	BPI-43	P19338 nucleolin (protein c23).	(R)SISLYYTGEK (K)NDLAVVDVR	8.66	48341
44	BPI-44	P07900 heat shock protein hsp 90-alpha	EDQTEYLEER	4.99	70465
45	BPI-45	P02533 keratin, type i cytoskeletal 14	(K)EVASNSELVQSSR (K)ASLEGNLAETENR (R)ISSVLAGGSCR (K)NHEEEMNALR (K)DAEEWFFTK	4.63	41917

46	BPI-46	P50991 t-complex protein 1, delta subunit		8.61	57516
47	BPI-47	P11142 heat shock cognate 71 kd protein	LLKDFFNGK CNEIINWLDK STAGDTHLGGEDFDNR (K)DAGTIAGLNVLR (R)IINEPTAAAIAYGLD K (R)TTPSYVAFTDTER (R)VEIIANDQGNR (K)EEFEHQQK (K)SQIHDIVLVGGSTR (R)FEELNADLFR	5.40	71346
48	BPI-48	P52895 probable trans-1,2-dihydrobenzene-1,2-diol dehydrogenase	(R)TPALIALR K)SIGVSNFNHR	7.85	37264
49	BPI-49	P02570 beta actin, cytoplasmic 1	(R)GYSFTTTAER	5.28	36758
50	BPI-50	P52907 f-actin capping protein alpha-1	(R)LLLNNDNLLR	5.44	34752
51	BPI-51	P07339 cathepsin d precursor (ec 3.4.23.5)	(R)YYTVFDR (K)LSPEDYTLK	5.27	31764

As will be evident to one of skill in the art, a given BPI can be described according to the data provided for that BPI in Table III or IV. The BPI is a protein comprising a peptide sequence described for that BPI (preferably comprising a plurality of, more preferably all of, the peptide sequences described for that BPI) and has a pI of about the value stated for that BPI (preferably within 10%, more preferably within 5% still more preferably within 1% of the stated value) and has a MW of about the value stated for that BPI (preferably within 10%, more preferably within 5%, still more preferably within 1% of the stated value).

In one embodiment, tissue from a subject is analyzed for quantitative detection of one or more of the following BPIs: BPI-1 to BPI-33 or any combination of them, wherein a decreased abundance of the BPI or BPIs (or any combination of them) in tissue from the subject relative to the same tissue from a subject or subjects free from breast cancer (e.g.,

a control sample or a previously determined reference range) indicates the presence of metastatic breast cancer.

In another embodiment of the invention, tissue from a subject is analyzed for quantitative detection of one or more of the following BPIs: BPI-34 to BPI-51, or any combination of them, wherein an increased abundance of the BPI or BPIs (or any combination of them) in tissue from the subject relative to the same tissue from a subject or subjects free from breast cancer (e.g., a control sample or a previously determined reference range) indicates the presence of breast cancer.

In another embodiment, tissue from a subject is analyzed for quantitative detection of one or more BPIs and one or more previously known biomarkers of breast cancer (e.g. CEA, CA-125). In accordance with this embodiment, the abundance of each BPI and known biomarkers relative to a control or reference range indicates whether a subject has metastatic breast cancer.

In a preferred embodiment, tissue from a subject is analyzed for quantitative detection of a cluster of BPIs.

As shown above, the BPIs described herein include isoforms of known proteins where the isoforms were not previously known to be associated with breast tissue. For each BPI, the present invention additionally provides: (a) a preparation comprising the isolated BPI; (b) a preparation comprising one or more fragments of the BPI; and (c) antibodies that bind to said BPI, to said fragments, or both to said BPI and to said fragments. As used herein, a BPI is "isolated" when it is present in a preparation that is substantially free of contaminating proteins, i.e., a preparation in which less than 10% (preferably less than 5%, more preferably less than 1%) of the total protein present is contaminating protein(s). A contaminating protein is a protein or protein isoform having a significantly different pI or MW from those of the isolated BPI, as determined by 2D electrophoresis. As used herein, a "significantly different" pI or MW is one that permits the contaminating protein to be resolved from the BPI on 2D electrophoresis, performed according to the Reference Protocol.

In one embodiment, an isolated protein is provided, said protein comprising a peptide with the amino acid sequence identified in Table III or IV for a BPI, said protein having a pI and MW within 10% (preferably within 5%, more preferably within 1%) of the values identified in Table III or IV for that BPI.

The BPIs of the invention can be assayed by any method known to those skilled in the art, including but not limited to the Preferred Technology described herein, enzyme assays, immunoassays, and western blotting. In one embodiment, the BPIs are separated on a 2-D gel by virtue of their MWs and pIs and visualized by staining the gel. In one embodiment, the BPIs are stained with a fluorescent dye and imaged with a fluorescence scanner. Sypro Red (Molecular Probes, Inc., Eugene, Oregon) is a suitable dye for this purpose. A preferred fluorescent dye is described in U.S. Application No. 09/412,168, filed on October 5, 1999, which is incorporated herein by reference in its entirety. Alternatively, BPIs can be detected in an immunoassay. In one embodiment, an immunoassay is performed by contacting a sample from a subject to be tested with an anti-BPI antibody under conditions such that immunospecific binding can occur if the BPI is present, and detecting or measuring the amount of any immunospecific binding by the antibody. Anti-BPI antibodies can be produced by the methods and techniques taught herein; examples of such antibodies known in the art are set forth in Table V. Preferably, the anti-BPI antibody preferentially binds to the BPI rather than to other isoforms of the same protein. In a preferred embodiment, the anti-BPI antibody binds to the BPI with at least 2-fold greater affinity, more preferably at least 5-fold greater affinity, still more preferably at least 10-fold greater affinity, than to said other isoforms of the same protein. Further information about these antibodies can be obtained from their commercial sources at:

Accurate Chemical & Scientific Corporation -http://www.accuratechemical.com

RDI Research Diagnostics, Inc. - http://www.researchd.com/

Table V Antibodies to BPIs of the present invention

BPI	Commercial Supplier	Cat. No.	Protein Name
BPI-1	Accurate Chemical Corp	YM- 3020	Cytokeratin 19
BPI-2	Accurate Chemical Corp	YSRT- MCA915G	Cytokeratin 18
BPI-3	Accurate Chemical Corp	MED- CLA 92/1	Cytokeratin 17
BPI-4	Accurate Chemical Corp	MED- CLA411	Cytokeratin 6
BPI-6	Research Diagnostics, Inc	RDI-CBL232	Cytokeratin 5, 8
BPI-7	Accurate Chemical Corp	MED- CLA 92/1	Cytokeratin 17
BPI-8	Accurate Chemical Corp	MED- CLA 92/1	Cytokeratin 17
BPI-9	Accurate Chemical Corp	YM- 3020	Cytokeratin 19

BPI-10	Accurate Chemical Corp	YM- 3020	Cytokeratin 19
BPI-11	Accurate Chemical Corp	MED- CLA411	Cytokeratin 6
BPI-12	Accurate Chemical Corp	MED- CLA411	Cytokeratin 6
BPI-13	Accurate Chemical Corp	MED- CLA 92/1	Cytokeratin 17
BPI-14	Accurate Chemical Corp	MED- CLA 92/1	Cytokeratin 17
BPI-15	Accurate Chemical Corp	MAB- 100	Creatine Kinase
BPI-17	Accurate Chemical Corp	YM- 3020	Cytokeratin 19
BPI-18	Accurate Chemical Corp	YM- 3020	Cytokeratin 19
BPI-19	Accurate Chemical Corp	MED- CLA 92/1	Cytokeratin 17
BPI-20	Accurate Chemical Corp	YM- 3020	Cytokeratin 19
BPI-21	Accurate Chemical Corp	MED- CLA 92/1	Cytokeratin 17
BPI-23	Accurate Chemical Corp	MAS- 445b	Cytokeratin 10
BPI-25	Accurate Chemical Corp	MAS- 445b	Cytokeratin 10
BPI-26	Accurate Chemical Corp	BMD- FM008	Cytokeratin 8
BPI-27	Accurate Chemical Corp	BMD- FM008	Cytokeratin 8
BPI-28	Accurate Chemical Corp	YM- 3014	Cytokeratin 7
BPI-28	Accurate Chemical Corp	YM- 3014	Cytokeratin 7
BPI-30	Accurate Chemical Corp	YSRT- MCA890	Cytokeratin 14
BPI-32	Accurate Chemical Corp	BYA- 9074-1	Cathepsin D 34kD
BPI-35	Accurate Chemical Corp	BMD- V3079	Tropomyosin, Muscle
BPI-36	Accurate Chemical Corp	YSRT- MCA358G	Actin
BPI-38	Accurate Chemical Corp	AXL- 163/2	Retinol Binding Protein
BPI-39	Research Diagnostics, Inc	RDI-PYRUKINabG	Pyruvate kinase
BPI-40	Accurate Chemical Corp	YSRT- MCA1538	Thioredoxin
BPI-42	Accurate Chemical Corp	MED- CLA 183	Heat Shock Protein 27
BPI-44	Accurate Chemical Corp	YSG- SPA835	Heat Shock Protein 90
BPI-45	Accurate Chemical Corp	YSRT- MCA890	Cytokeratin 14
BPI-47	Accurate Chemical Corp	YSG- SPA820AP	Heat Shock Protein 70
BPI-51	Accurate Chemical Corp	BYA- 9074-1	Cathepsin D 34kD

In one embodiment, binding of antibody in tissue sections can be used to detect aberrant BPI localization or an aberrant level of one or more BPIs. In a specific embodiment, antibody to a BPI can be used to assay a patient tissue (e.g., a breast biopsy) for the level of the BPI where an aberrant level of BPI is indicative of breast cancer. As used herein, an "aberrant level" means a level that is increased or decreased compared with the level in a subject free from breast cancer or a reference level. If desired, the comparison can be performed with a matched sample from the same subject, taken from a portion of the body not affected by breast cancer.

Suitable immunoassays include, without limitation, competitive and non-competitive assay systems using techniques such as western blots, radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoprecipitation

assays, precipitin reactions, gel diffusion precipitin reactions, immunodiffusion assays, agglutination assays, complement-fixation assays, immunoradiometric assays, fluorescent immunoassays and protein A immunoassays.

If desired, a BPI can be detected by means of a two-step sandwich assay. In the first step, a capture reagent (e.g., an anti-BPI antibody) is used to capture the BPI. The capture reagent can optionally be immobilized on a solid phase. In the second step, a directly or indirectly labeled detection reagent is used to detect the captured BPI. In one embodiment, the detection reagent is a lectin. Any lectin can be used for this purpose that preferentially binds to the BPI rather than to other isoforms that have the same core protein as the BPI or to other proteins that share the antigenic determinant recognized by the antibody. In a preferred embodiment, the chosen lectin binds to the BPI with at least 2-fold greater affinity, more preferably at least 5-fold greater affinity, still more preferably at least 10-fold greater affinity, than to said other isoforms that have the same core protein as the BPI or to said other proteins that share the antigenic determinant recognized by the antibody. A lectin that is suitable for detecting a given BPI can readily be identified by methods well known in the art, for instance upon testing one or more lectins enumerated in Table I on pages 158-159 of Sumar et al., Lectins as Indicators of Disease-Associated Glycoforms, In: Gabius H-J & Gabius S (eds.), 1993, Lectins and Glycobiology, at pp. 158-174 (which is incorporated herein by reference in its entirety).

In an alternative embodiment, the detection reagent is an antibody, e.g., an antibody that immunospecifically detects other post-translational modifications, such as an antibody that immunospecifically binds to phosphorylated amino acids. Examples of such antibodies include those that bind to phosphotyrosine (BD Transduction Laboratories, catalog nos.: P11230-050/P11230-150; P11120; P38820; P39020), those that bind to phosphoserine (Zymed Laboratories Inc., catalog no. 61-8100) and those that bind to phosphothreonine (Zymed Laboratories Inc., catalog nos. 71-8200, 13-9200).

If desired, a gene encoding a BPI, a related gene, or related nucleic acid sequences or subsequences, including complementary sequences, can also be used in hybridization assays. A nucleotide encoding a BPI, or subsequences thereof comprising at least 8 nucleotides, can be used as a hybridization probe. Hybridization assays can be used for detection, prognosis, diagnosis, or monitoring of conditions, disorders, or disease states, associated with aberrant expression of genes encoding BPIs, or for differential diagnosis of patients with signs or symptoms suggestive of breast cancer. In particular, such a

hybridization assay can be carried out by a method comprising contacting a patient sample containing nucleic acid with a nucleic acid probe capable of hybridizing to a DNA or RNA that encodes a BPI, under conditions such that hybridization can occur, and detecting or measuring any resulting hybridization. Nucleotides can be used for therapy of patients having breast cancer, as described below.

The invention also provides diagnostic kits, comprising an anti-BPI antibody. In addition, such a kit may optionally comprise one or more of the following: (1) instructions for using the anti-BPI antibody for diagnosis, prognosis, therapeutic monitoring or any combination of these applications; (2) a labeled binding partner to the antibody; (3) a solid phase (such as a reagent strip) upon which the anti-BPI antibody is immobilized; and (4) a label or insert indicating regulatory approval for diagnostic, prognostic or therapeutic use or any combination thereof. If no labeled binding partner to the antibody is provided, the anti-BPI antibody itself can be labeled with a detectable marker, e.g., a chemiluminescent, enzymatic, fluorescent, or radioactive moiety.

The invention also provides a kit comprising a nucleic acid probe capable of hybridizing to RNA encoding a BPI. In a specific embodiment, a kit comprises in one or more containers a pair of primers (e.g., each in the size range of 6-30 nucleotides, more preferably 10-30 nucleotides and still more preferably 10-20 nucleotides) that under appropriate reaction conditions can prime amplification of at least a portion of a nucleic acid encoding a BPI, such as by polymerase chain reaction (see e.g., Innis et al., 1990, PCR Protocols, Academic Press, Inc., San Diego, CA), ligase chain reaction (see EP 320,308) use of Qb replicase, cyclic probe reaction, or other methods known in the art. Kits are also provided which allow for the detection of a plurality of BPIs or a plurality of nucleic acids each encoding a BPI. A kit can optionally further comprise a predetermined amount of an isolated BPI protein or a nucleic acid encoding a BPI, e.g., for use as a standard or control.

4.3. Techniques for Identifying BFs/BPIs and BF/BPI Clusters

In most cases, the metastatic breast cancer disease process is likely to be associated with a combination of BFs or BPIs (and to be regulated by a combination of BPIs), rather than individual BFs and BPIs in isolation. The strategies for discovering such combinations of BFs and BPIs differ from those for discovering individual BFs and BPIs. In such cases,

each individual BF and BPI can be regarded as one variable and the disease can be regarded as a joint, multi-variate effect caused by interaction of these variables.
The following steps can be used to identify BFs and BPIs from data produced by the Preferred Technology.
The first step is to identify a collection of BFs or BPIs that individually show significant association with breast luminal epithelial cells when compared with myoepithelial cells. The association between the identified BFs or BPIs and metastatic breast cancer need not be as highly significant as is desirable when an individual BF or BPI is used as a diagnostic. Any of a number of statistical tests (fold changes, Wilcoxon Rank Sum test, etc.) can be used. Once a suitable collection of BFs or BPIs has been identified, a sophisticated multi-variate analysis capable of identifying clusters can then be used to estimate the significant multivariate associations with luminal versus myoepithelial breast cells.
Linear Discriminant Analysis (LDA) is one such procedure, which can be used to detect significant association between a cluster of variables (i.e., BFs or BPIs) between luminal and myoepithelial breast cells.

4.4. Use in Clinical Studies

The diagnostic methods and compositions of the present invention can assist in monitoring a clinical study, e.g. to evaluate drugs for therapy of breast cancer. In one embodiment, candidate molecules are tested for their ability to restore BF or BPI levels in a patient having breast cancer to levels found in subjects free from breast cancer or, in a treated patient (e.g. after treatment with a chemotherapeutic agent), to preserve BF or BPI levels at or near non-breast cancer values. The levels of one or more BFs or BPIs can be assayed.
In another embodiment, the methods and compositions of the present invention are used to screen candidates for a clinical study to identify individuals having metastatic breast cancer; such individuals can then be excluded from the study or can be placed in a separate cohort for treatment or analysis

4.4.1 Purification of BPIs

In particular aspects, the invention provides isolated BPIs, preferably human BPIs, and fragments and derivatives thereof which comprise an antigenic determinant (i.e., can be

recognized by an antibody) or which are otherwise functionally active, as well as nucleic acid sequences encoding the foregoing. "Functionally active" as used herein refers to material displaying one or more functional activities associated with a full-length (wild-type) BPI, e.g., binding to a BPI substrate or BPI binding partner, antigenicity (binding to an anti-target antibody), immunogenicity, enzymatic activity etc.

In specific embodiments, the invention provides fragments of a BPI comprising at least 5 amino acids, at least 10 amino acids, at least 50 amino acids, or at least 75 amino acids. Fragments lacking some or all of the regions of a BPI are also provided, as are proteins (e.g., fusion proteins) comprising such fragments. Nucleic acids encoding the foregoing are provided.

Once a recombinant nucleic acid which encodes the BPI, a portion of the BPI, or a precursor of the BPI is identified, the gene product can be analyzed. This is achieved by assays based on the physical or functional properties of the product, including radioactive labeling of the product followed by analysis by gel electrophoresis, immunoassay, etc. The BPIs identified herein can be isolated and purified by standard methods including chromatography (e.g., ion exchange, affinity, and sizing column chromatography), centrifugation, differential solubility, or by any other standard technique for the purification of proteins.

Alternatively, once a recombinant nucleic acid that encodes the BPI is identified, the entire amino acid sequence of the BPI can be deduced from the nucleotide sequence of the gene coding region contained in the recombinant nucleic acid. As a result, the protein can be synthesized by standard chemical methods known in the art (e.g., see Hunkapiller et al., 1984, Nature 310:105-111).

In another alternative embodiment, native BPIs can be purified from natural sources, by standard methods such as those described above (e.g., immunoaffinity purification).

In a preferred embodiment, BPIs are isolated by the Preferred Technology described in U.S. Application No. 08/980,574, which is incorporated herein by reference. For preparative-scale runs, a narrow-range "zoom gel" having a pH range of 2 pH units or less is preferred for the isoelectric step, according to the method described in Westermeier, 1993, Electrophoresis in Practice (VCH, Weinheim, Germany), pp. 197-209 (which is incorporated herein by reference in its entirety); this modification permits a larger quantity of a target protein to be loaded onto the gel, and thereby increases the quantity of is lated BPI that can be recovered from the gel. When used in this way for preparative-

scale runs, the Preferred Technology typically provides up to 100 ng, and can provide up to 1000 ng, of an isolated BPI in a single run. Those of skill in the art will appreciate that a zoom gel can be used in any separation strategy which employs gel isoelectric focusing. The invention thus encompasses a BPI, a fragment or derivative of a BPI, and homologs and analogs of a BPI; any of the foregoing can be produced by recombinant DNA techniques or by chemical synthetic methods.

4.5. Isolation Of DNA Encoding A BPI

Specific embodiments for the cloning of a gene encoding a BPI, are presented below by way of example and not of limitation.

The nucleotide sequences of the present invention, including DNA and RNA, and comprising a sequence encoding the BPI (or a fragment, homologue or analog thereof), may be synthesized using methods known in the art, such as using conventional chemical approaches or polymerase chain reaction (PCR) amplification. The nucleotide sequences of the present invention also permit the identification and cloning of the gene encoding a BPI from any species, for instance by screening cDNA libraries, genomic libraries or expression libraries.

For example, to clone a gene encoding a BPI by PCR techniques, anchored degenerate oligonucleotides (or a set of most likely oligonucleotides) can be designed for all BPI peptide fragments identified as part of the same protein. PCR reactions under a variety of conditions can be performed with relevant cDNA and genomic DNAs (e.g., from breast tissue) from one or more species. Also vectorette reactions can be performed on any available cDNA and genomic DNA using the oligonucleotides (which preferably are nested) as above. Vectorette PCR is a method that enables the amplification of specific DNA fragments in situations where the sequence of only one primer is known. Thus, it extends the application of PCR to stretches of DNA where the sequence information is only available at one end. (Arnold C, 1991, PCR Methods Appl. 1(1):39-42; Dyer KD, Biotechniques, 1995, 19(4):550-2). Vectorette PCR may pe performed with probes that are anchored degenerate oligonucleotides (or most likely oligonucleotides) coding for BPI peptide fragments, using as a template a genomic library or cDNA library pools. Anchored degenerate and most likely oligonucleotides can be designed for all BPI peptide fragments. These oligonucleotides may be labelled and hybridized to filters containing cDNA and genomic DNA libraries. Oligonucleotides to different peptides

from the same protein will often identify the same members of the library. The cDNA and genomic DNA libraries may be obtained from multiple mammalian species. Nucleotide sequences comprising a nucleotide sequence encoding a BPI of the present invention are useful for their ability to hybridize selectively with complementary stretches of genes encoding other proteins. Depending on the application, a variety of hybridization conditions may be employed to obtain nucleotide sequences at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99%, or 100% identical to the sequence of a nucleotide encoding a BPI.

For a high degree of selectivity, relatively stringent conditions are used to form the duplexes, such as low salt or high temperature conditions. As used herein, "highly stringent conditions" means hybridization to filter-bound DNA in 0.5 M NaHPO4, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1xSSC/0.1% SDS at 68°C (Ausubel F.M. et al., eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & Sons, Inc., New York, at p. 2.10.3; incorporated herein by reference in its entirety.) For some applications, less stringent conditions for duplex formation are required. As used herein "moderately stringent conditions" means washing in 0.2xSSC/0.1% SDS at 42°C (Ausubel et al., 1989, supra). Hybridization conditions can also be rendered more stringent by the addition of increasing amounts of formamide, to destabilize the hybrid duplex. Thus, particular hybridization conditions can be readily manipulated, and will generally be chosen depending on the desired results. In general, convenient hybridization temperatures in the presence of 50% formamide are: 42°C for a probe which is 95 to 100% identical to the fragment of a gene encoding a BPI, 37°C for 90 to 95% identity and 32°C for 70 to 90% identity.

In the preparation of genomic libraries, DNA fragments are generated, some of which will encode parts or the whole of a BPI. The DNA may be cleaved at specific sites using various restriction enzymes. Alternatively, one may use DNAse in the presence of manganese to fragment the DNA, or the DNA can be physically sheared, as for example, by sonication. The DNA fragments can then be separated according to size by standard techniques, including but not limited to agarose and polyacrylamide gel electrophoresis, column chromatography and sucrose gradient centrifugation. The DNA fragments can then be inserted into suitable vectors, including but not limited to plasmids, cosmids, bacteriophages lambda or T4, and yeast artificial chromosome (YAC). (See, for example,

Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; Glover, D.M. (ed.), 1985, DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, U.K. Vol. I, II; Ausubel F.M. et al., eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York). The genomic library may be screened by nucleic acid hybridization to labeled probe (Benton and Davis, 1977, Science 196:180; Grunstein and Hogness, 1975, Proc. Natl. Acad. Sci. U.S.A. 72:3961).
The genomic libraries may be screened with labeled degenerate oligonucleotide probes corresponding to the amino acid sequence of any peptide of the BPI using optimal approaches well known in the art. Any probe used preferably is 10 nucleotides or longer, more preferably 15 nucleotides or longer.
As shown in Tables III and IV above, some BPIs disclosed herein correspond to isoforms of previously identified proteins encoded by genes whose sequences are publicly known. To screen such a gene, any probe may be used that is complementary to the gene or its complement; preferably the probe is 10 nucleotides or longer, more preferably 15 nucleotides or longer. The Entrez database held by the National Center for Biotechnology Information (NCBI) - which is accessible at http://www.ncbi.nlm.nih.gov/ - provides gene sequences for these BPIs under the accession numbers detailed in the third column of Tables III and IV, and each sequence is incorporated herein by reference

In yet another aspect of the invention, clones containing nucleotide sequences encoding the entire BPI or a part thereof, or a BPI-derived polypeptide may also be obtained by screening expression libraries. For example, DNA from the relevant source is isolated and random fragments are prepared and ligated into an expression vector (e.g., a bacteriophage, plasmid, phagemid or cosmid) such that the inserted sequence in the vector is capable of being expressed by the host cell into which the vector is then introduced. Various screening assays can then be used to select for the expressed BPI or BPI-derived polypeptides. In one embodiment, the various anti-BPI antibodies of the invention can be used to identify the desired clones using methods known in the art. See, for example, Harlow and Lane, 1988, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, Appendix IV. Colonies or plaques from the library are brought into contact with the antibodies to identify those clones that bind antibody.

In an embodiment, colonies or plaques containing DNA that encodes a BPI or BPI-derived polypeptide can be detected using DYNA Beads according to Olsvick et al., 29th ICAAC, Houston, Tex. 1989, incorporated herein by reference. Anti-BPI antibodies are crosslinked to tosylated DYNA Beads M280, and these antibody-containing beads are then contacted with colonies or plaques expressing recombinant polypeptides. Colonies or plaques expressing a BPI or BPI-derived polypeptide are identified as any of those that bind the beads.

Alternatively, the anti-BPI antibodies can be nonspecifically immobilized to a suitable support, such as silica or Celite resin. This material is then used to adsorb to bacterial colonies expressing the BPI protein or BPI-derived polypeptide as described herein.

In another aspect, PCR amplification may be used to isolate from genomic DNA a substantially pure DNA (i.e., a DNA substantially free of contaminating nucleic acids) encoding the entire BPI or a part thereof. Preferably such a DNA is at least 95% pure, more preferably at least 99% pure. Oligonucleotide sequences, degenerate or otherwise, corresponding to known BPI sequences can be used as primers.

PCR can be carried out, e.g., by use of a Perkin-Elmer thermal cycler (e.g. GeneAmp9700) and Taq polymerase (e.g. AmpliTaq DNA polymerase). One can choose to synthesize several different degenerate primers, for use in the PCR reactions. It is also possible to vary the stringency of hybridization conditions used in priming the PCR reactions, to allow for greater or lesser degrees of nucleotide sequence similarity between the degenerate primers and the corresponding sequences in the DNA. After successful amplification of a segment of the sequence encoding a BPI, that segment may be molecularly cloned and sequenced, and utilized as a probe to isolate a complete genomic clone. This, in turn, will permit the determination of the gene's complete nucleotide sequence, the analysis of its expression, and the production of its protein product for functional analysis, as described infra.

The gene encoding a BPI can also be identified by mRNA selection by nucleic acid hybridization followed by in vitro translation. In this procedure, fragments are used to isolate complementary mRNAs by hybridization. Such DNA fragments may represent available, purified DNA encoding a BPI of another species (e.g., mouse, human). Immunoprecipitation analysis or functional assays (e.g., aggregation ability in vitro; binding to receptor) of the in vitro translation products of the isolated products of the isolated mRNAs identifies the mRNA and, therefore, the complementary DNA fragments

that contain the desired sequences. In addition, specific mRNAs may be selected by adsorption of polysomes isolated from cells to immobilized antibodies that specifically recognize a BPI. A radiolabelled cDNA encoding a BPI can be synthesized using the selected mRNA (from the adsorbed polysomes) as a template. The radiolabelled mRNA or cDNA may then be used as a probe to identify the DNA fragments encoding a BPI from among other genomic DNA fragments.

Alternatives to isolating genomic DNA encoding a BPI include, but are not limited to, chemically synthesizing the gene sequence itself from a known sequence or making cDNA to the mRNA which encodes the BPI. For example, RNA for cDNA cloning of the gene encoding a BPI can be isolated from cells which express the BPI. Other methods are possible and within the scope of the invention.

Any eukaryotic cell can serve as the nucleic acid source for the molecular cloning of the gene encoding a BPI. The nucleic acid sequences encoding the BPI can be isolated from vertebrate, mammalian, primate, human, porcine, bovine, feline, avian, equine, canine or murine sources. The DNA may be obtained by standard procedures known in the art from cloned DNA (e.g., a DNA "library"), by chemical synthesis, by cDNA cloning, or by the cloning of genomic DNA, or fragments thereof, purified from the desired cell. (See, for example, Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; Glover, D.M. (ed.), 1985, DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, U.K. Vol. I, II.) Clones derived from genomic DNA may contain regulatory and intron DNA regions in addition to coding regions; clones derived from cDNA will contain only exon sequences.

The identified and isolated gene or cDNA can then be inserted into an appropriate cloning vector. A large number of vector-host systems known in the art may be used. The only limitation is that the vector system chosen be compatible with the host cell used. Such vectors include, but are not limited to, bacteriophages such as lambda derivatives, plasmids such as PBR322 or pUC plasmid derivatives or the Bluescript vector (Stratagene) or modified viruses such as adenoviruses, adeno-associated viruses or retroviruses. The insertion into a cloning vector can, for example, be accomplished by ligating the DNA fragment into a cloning vector which has complementary cohesive termini. However, if the complementary restriction sites used to fragment the DNA are not present in the cloning vector, the ends of the DNA molecules may be enzymatically

modified. Alternatively, any site desired may be produced by ligating nucleotide sequences (linkers) onto the DNA termini; these ligated linkers may comprise specific chemically synthesized oligonucleotides encoding restriction endonuclease recognition sequences. In an alternative method, the cleaved vector and the gene encoding a BPI may be modified by homopolymeric tailing. Recombinant molecules can be introduced into host cells via transformation, transfection, infection, electroporation, etc., so that many copies of the gene sequence are generated.

In specific embodiments, transformation of host cells with recombinant DNA molecules that incorporate the isolated gene encoding the BPI, cDNA, or synthesized DNA sequence enables generation of multiple copies of the gene. Thus, the gene may be obtained in large quantities by growing transformants, isolating the recombinant DNA molecules from the transformants and, when necessary, retrieving the inserted gene from the isolated recombinant DNA.

The nucleotide sequences of the present invention include nucleotide sequences encoding amino acid sequences with substantially the same amino acid sequences as native BPIs, and nucleotide sequences encoding amino acid sequences with functionally equivalent amino acids, as well as those encoding other target derivatives or analogs.

4.6. Expression of DNA Encoding BPIs

The nucleotide sequence coding for a BPI or a functionally active analog or fragment or other derivative thereof can be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted protein-coding sequence. The necessary transcriptional and translational signals can also be supplied by the native gene encoding the BPI or its flanking regions. A variety of host-vector systems may be utilized to express the protein-coding sequence. These include but are not limited to mammalian cell systems infected with virus (e.g., vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (e.g., baculovirus); microorganisms such as yeast containing yeast vectors; or bacteria transformed with bacteriophage, DNA, plasmid DNA, or cosmid DNA. The expression elements of vectors vary in their strengths and specificities. Depending on the host-vector system utilized, any one of a number of suitable transcription and translation elements may be used. In specific embodiments, a nucleotide sequence encoding a human gene (or a nucleotide sequence encoding a functionally active portion of a human

BPI) is expressed. In yet another embodiment, a fragment of a BPI comprising a domain of the BPI is expressed.

Any of the methods previously described for the insertion of DNA fragments into a vector may be used to construct expression vectors containing a chimeric gene consisting of appropriate transcriptional and translational control signals and the protein coding sequences. These methods may include in vitro recombinant DNA and synthetic techniques and in vivo recombinants (genetic recombination). Expression of nucleic acid sequence encoding a BPI or fragment thereof may be regulated by a second nucleic acid sequence so that the BPI or fragment is expressed in a host transformed with the recombinant DNA molecule. For example, expression of a BPI may be controlled by any promoter or enhancer element known in the art. Promoters which may be used to control the expression of the gene encoding a BPI include, but are not limited to, the SV40 early promoter region (Bernoist and Chambon, 1981, Nature 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto, et al., 1980, Cell 22:787-797), the herpes thymidine kinase promoter (Wagner et al., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445), the regulatory sequences of the metallothionein gene (Brinster et al., 1982, Nature 296:39-42), the tetracycline (Tet) promoter (Gossen et al., 1995, Proc. Nat. Acad. Sci. USA 89:5547-5551); prokaryotic expression vectors such as the beta-lactamase promoter (Villa-Kamaroff, et al., 1978, Proc. Natl. Acad. Sci. U.S.A. 75:3727-3731), or the tac promoter (DeBoer, et al., 1983, Proc. Natl. Acad. Sci. U.S.A. 80:21-25; see also "Useful proteins from recombinant bacteria" in Scientific American, 1980, 242:74-94); plant expression vectors comprising the nopaline synthetase promoter region (Herrera-Estrella et al., Nature 303:209-213) or the cauliflower mosaic virus 35S RNA promoter (Gardner, et al., 1981, Nucl. Acids Res. 9:2871), and the promoter of the photosynthetic enzyme ribulose biphosphate carboxylase (Herrera-Estrella et al., 1984, Nature 310:115-120); promoter elements from yeast or other fungi such as the Gal 4 promoter, the ADC (alcohol dehydrogenase) promoter, PGK (phosphoglycerol kinase) promoter, alkaline phosphatase promoter, and the following animal transcriptional control regions, which exhibit tissue specificity and have been utilized in transgenic animals: elastase I gene control region which is active in pancreatic acinar cells (Swift et al., 1984, Cell 38:639-646; Ornitz et al., 1986, Cold Spring Harbor Symp. Quant. Biol. 50:399-409; MacDonald, 1987, Hepatology 7:425-515); insulin gene control region which is active in pancreatic beta cells (Hanahan, 1985, Nature 315:115-122), immunoglobulin gene control

region which is active in lymphoid cells (Grosschedl et al., 1984, Cell 38:647-658; Adames et al., 1985, Nature 318:533-538; Alexander et al., 1987, Mol. Cell. Biol. 7:1436-1444), mouse mammary tumor virus control region which is active in testicular, breast, lymphoid and mast cells (Leder et al., 1986, Cell 45:485-495), Whey Acidic Protein promoter which has been used to induce mammary gland-specific expression (Sympson CJ et al., 1994 J Cell Biol 125: 681-93)albumin gene control region which is active in liver (Pinkert et al., 1987, Genes and Devel. 1:268-276), alpha-fetoprotein gene control region which is active in liver (Krumlauf et al., 1985, Mol. Cell. Biol. 5:1639-1648; Hammer et al., 1987, Science 235:53-58; alpha 1-antitrypsin gene control region which is active in the liver (Kelsey et al., 1987, Genes and Devel. 1:161-171), beta-globin gene control region which is active in myeloid cells (Mogram et al., 1985, Nature 315:338-340; Kollias et al., 1986, Cell 46:89-94; myelin basic protein gene control region which is active in oligodendrocyte cells in the breast (Readhead et al., 1987, Cell 48:703-712); myosin light chain-2 gene control region which is active in skeletal muscle (Sani, 1985, Nature 314:283-286); neuronal-specific enolase (NSE) which is active in neuronal cells (Morelli et al., 1999, Gen. Virol. 80:571-83); breast-derived neurotrophic factor (BDNF) gene control region which is active in neuronal cells (Tabuchi et al., 1998, Biochem. Biophysic. Res. Com. 253:818-823); glial fibrillary acidic protein (GFAP) promoter which is active in astrocytes (Gomes et al., 1999, Braz J Med Biol Res 32(5):619-631; Morelli et al., 1999, Gen. Virol. 80:571-83) and gonadotropic releasing hormone gene control region which is active in the hypothalamus (Mason et al., 1986, Science 234:1372-1378).

In a specific embodiment, a vector is used that comprises a promoter operably linked to a BPI-encoding nucleic acid, one or more origins of replication, and, optionally, one or more selectable markers (e.g., an antibiotic resistance gene).

In a specific embodiment, an expression construct is made by subcloning a BPI coding sequence into the EcoRI restriction site of each of the three pGEX vectors (Glutathione S-Transferase expression vectors; Smith and Johnson, 1988, Gene 7:31-40). This allows for the expression of the BPI product from the subclone in the correct reading frame.

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the BPI coding sequence may be ligated to an adenovirus transcription/translation control complex, e.g., the late

promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing the antibody molecule in infected hosts. (e.g., see Logan & Shenk, 1984, Proc. Natl. Acad. Sci. USA 81:355-359). Specific initiation signals may also be required for efficient translation of inserted antibody coding sequences. These signals include the ATG initiation codon and adjacent sequences. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner et al., 1987, Methods in Enzymol. 153:51-544).

Expression vectors containing inserts of a gene encoding a BPI can be identified by three general approaches: (a) nucleic acid hybridization, (b) presence or absence of "marker" gene functions, and (c) expression of inserted sequences. In the first approach, the presence of a gene encoding a BPI inserted in an expression vector can be detected by nucleic acid hybridization using probes comprising sequences that are homologous to an inserted gene encoding a BPI. In the second approach, the recombinant vector/host system can be identified and selected based upon the presence or absence of certain "marker" gene functions (e.g., thymidine kinase activity, resistance to antibiotics, transformation phenotype, occlusion body formation in baculovirus, etc.) caused by the insertion of a gene encoding a BPI in the vector. For example, if the gene encoding the BPI is inserted within the marker gene sequence of the vector, recombinants containing the gene encoding the BPI insert can be identified by the absence of the marker gene function. In the third approach, recombinant expression vectors can be identified by assaying the gene product (i.e., BPI) expressed by the recombinant. Such assays can be based, for example, on the physical or functional properties of the BPI in in vitro assay systems, e.g., binding with anti-BPI antibody.

In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Expression from certain promoters can be elevated in the presence of certain inducers; thus, expression of the genetically engineered BPI may be controlled.

Furthermore, different host cells have characteristic and specific mechanisms for the translational and post-translational processing and modification (e.g., glycosylation, phosphorylation of proteins). Appropriate cell lines or host systems can be chosen to ensure the desired modification and processing of the foreign protein expressed. For example, expression in a bacterial system will produce an unglycosylated product and expression in yeast will produce a glycosylated product. Eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERY, BHK, Hela, COS, MDCK, 293, 3T3, WI38, and in particular, breast cancer cell lines such as, for example, BT483, Hs578T, HTB2, BT20 and T47D, and normal mammary gland cell lines such as, for example, CRL7030 and Hs578Bst. Furthermore, different vector/host expression systems may effect processing reactions to different extents.

For long-term, high-yield production of recombi-nant proteins, stable expression is preferred. For example, cell lines which stably express the differentially expressed or pathway gene protein may be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription termina-tors, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched medium, and then are switched to a selective medium. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which express the differentially expressed or pathway gene protein. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the differentially expressed or pathway gene protein.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, et al., 1977, Cell 11:223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, 1962, Proc. Natl. Acad. Sci. USA 48:2026), and adenine phosphoribosyltransferase (Lowy, et al., 1980, Cell 22:817) genes can be employed in tk-, hgprt- or aprt- cells, respectively. Also, antimetabolite resistance

can be used as the basis of selection for dhfr, which confers resistance to methotrexate (Wigler, et al., 1980, Natl. Acad. Sci. USA 77:3567; O'Hare, et al., 1981, Proc. Natl. Acad. Sci. USA 78:1527); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, 1981, Proc. Natl. Acad. Sci. USA 78:2072); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, et al., 1981, J. Mol. Biol. 150:1); and hygro, which confers resistance to hygromycin (Santerre, et al., 1984, Gene 30:147) genes.

In other specific embodiments, the BPI, fragment, analog, or derivative may be expressed as a fusion, or chimeric protein product (comprising the protein, fragment, analog, or derivative joined via a peptide bond to a heterologous protein sequence). For example, the polypeptides of the present invention may be fused with the constant domain of immunoglobulins (IgA, IgE, IgG, IgM), or portions thereof (CH1, CH2, CH3, or any combination thereof and portions thereof) resulting in chimeric polypeptides. Such fusion proteins may facilitate purification, increase half-life in vivo, and enhance the delivery of an antigen across an epithelial barrier to the immune system. An increase in the half-life in vivo and facilitated purification has been shown for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. See, e.g., EP 394,827; Traunecker et al., Nature, 331:84-86 (1988). Enhanced delivery of an antigen across the epithelial barrier to the immune system has been demonstrated for antigens (e.g., insulin) conjugated to an FcRn binding partner such as IgG or Fc fragments (see, e.g., PCT publications WO 96/22024 and WO 99/04813).

Nucleic acids encoding a BPI can fused to an epitope tag (e.g., the hemagglutinin ("HA") tag or flag tag) to aid in detection and purification of the expressed polypeptide. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht et al., 1991, Proc. Natl. Acad. Sci. USA 88:8972-897).

Fusion proteins can be made by ligating the appropriate nucleic acid sequences encoding the desired amino acid sequences to each other by methods known in the art, in the proper coding frame, and expressing the chimeric product by methods commonly known in the art. Alternatively, a fusion protein may be made by protein synthetic techniques, e.g., by use of a peptide synthesizer.

Both cDNA and genomic sequences can be cloned and expressed.

4.7. Production of Antibodies to BPIs

According to the invention a BPI, its fragments or other derivatives, or analogs thereof, may be used as an immunogen to generate antibodies which immunospecifically bind such an immunogen. Such proteins, fragments, derivatives or analogs can be isolated by any convenient means, including the methods described above. Antibodies of the invention include, but are not limited to polyclonal, monoclonal, bispecific, humanized or chimeric antibodies, single chain antibodies, Fab fragments and F(ab') fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site which specifically binds an antigen. The immunoglobulin molecules of the invention can be of any class (e.g., IgG, IgE, IgM, IgD and IgA) or subclass of immunoglobulin molecule.

In one embodiment, antibodies against a BPI are publicly available. Further information about these antibodies can be obtained from commercial sources such as:

Accurate Chemical & Scientific Corporation -http://www.accuratechemical.com
RDI Research Diagnostics, Inc - http://www.researchd.com/

In another embodiment, antibodies capable of binding immunospecifically a BPI or BPI fragment, but no other protein product of the gene from which the BPI is translated, nor any protein product of any other gene family member, are produced by methods known to those of skill in the art.

In one embodiment of the invention, antibodies to a specific domain of a BPI are produced. In a specific embodiment, hydrophilic fragments of a BPI are used as immunogens for antibody production.

In the production of antibodies, screening for the desired antibody can be accomplished by techniques known in the art, e.g. ELISA (enzyme-linked immunosorbent assay). For example, to select antibodies which recognize a specific domain of a BPI, one may assay generated hybridomas for a product which binds to a BPI fragment containing such domain. For selection of an antibody that specifically binds a first BPI homolog but which does not specifically bind to (or binds less avidly to) a second BPI homolog, one can select on the basis of positive binding to the first BPI homolog and a lack of binding

to (or reduced binding to) the second BPI homolog. Similarly, for selection of an antibody that specifically binds a BPI but which does not specifically bind to (or binds less avidly to) a different isoform of the same protein (such as a different glycoform having the same core peptide as the BPI), one can select on the basis of positive binding to the BPI and a lack of binding to (or reduced binding to) the different isoform (e.g., a different glycoform). Thus, the present invention provides an antibody (preferably a monoclonal antibody) that binds with greater affinity (preferably at least 2-fold, more preferably at least 5-fold still more preferably at least 10-fold greater affinity) to a BPI than to a different isoform or isoforms (e.g., glycoforms) of the BPI.

Polyclonal antibodies which may be used in the methods of the invention are heterogeneous populations of antibody molecules derived from the sera of immunized animals. Unfractionated immune serum can also be used. Various procedures known in the art may be used for the production of polyclonal antibodies to a BPI or derivative or analog. In a particular embodiment, rabbit polyclonal antibodies to an epitope of a BPI can be obtained. For example, for the production of polyclonal or monoclonal antibodies, various host animals can be immunized by injection with the native BPI, or a synthetic (e.g., recombinant) version, or derivative (e.g., fragment) thereof, including but not limited to rabbits, mice, rats, etc. The Preferred Technology described herein provides isolated BPIs suitable for such immunization. If the BPI is purified by gel electrophoresis, the BPI can be used for immunization with or without prior extraction from the polyacrylamide gel. Various adjuvants may be used to enhance the immunological response, depending on the host species, including, but not limited to, complete or incomplete Freund's adjuvant, a mineral gel such as aluminum hydroxide, surface active substance such as lysolecithin, pluronic polyol, a polyanion, a peptide, an oil emulsion, keyhole limpet hemocyanin, dinitrophenol, and an adjuvant such as BCG (bacille Calmette-Guerin) or corynebacterium parvum. Additional adjuvants are also well known in the art.

For preparation of monoclonal antibodies (mAbs) directed toward a BPI or fragment or analog thereof, any technique which provides for the production of antibody molecules by continuous cell lines in culture may be used. For example, the hybridoma technique originally developed by Kohler and Milstein (1975, Nature 256:495-497), as well as the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and the EBV-hybridoma technique to produce human

monoclonal antibodies (Cole et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAbs of the invention may be cultivated in vitro or in vivo. In an additional embodiment of the invention, monoclonal antibodies can be produced in germ-free animals utilizing known technology (PCT/US90/02545, incorporated herein by reference).

The monoclonal antibodies include but are not limited to human monoclonal antibodies and chimeric monoclonal antibodies (e.g., human-mouse chimeras). A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a human immunoglobulin constant region and a variable region derived from a murine mAb. (See, e.g., Cabilly et al., U.S. Patent No. 4,816,567; and Boss et al., U.S. Patent No. 4,816397, which are incorporated herein by reference in their entirety.) Humanized antibodies are antibody molecules from non-human species having one or more complementarily determining regions (CDRs) from the non-human species and a framework region from a human immunoglobulin molecule. (See, e.g., Queen, U.S. Patent No. 5,585,089, which is incorporated herein by reference in its entirety.) Chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT Publication No. WO 87/02671; European Patent Application 184,187; European Patent Application 171,496; European Patent Application 173,494; PCT Publication No. WO 86/01533; U.S. Patent No. 4,816,567; European Patent Application 125,023; Better et al., 1988, Science 240:1041-1043; Liu et al., 1987, Proc. Natl. Acad. Sci. USA 84:3439-3443; Liu et al., 1987, J. Immunol. 139:3521-3526; Sun et al., 1987, Proc. Natl. Acad. Sci. USA 84:214-218; Nishimura et al., 1987, Canc. Res. 47:999-1005; Wood et al., 1985, Nature 314:446-449; and Shaw et al., 1988, J. Natl. Cancer Inst. 80:1553-1559; Morrison, 1985, Science 229:1202-1207; Oi et al., 1986, Bio/Techniques 4:214; U.S. Patent 5,225,539; Jones et al., 1986, Nature 321:552-525; Verhoeyan et al. (1988) Science 239:1534; and Beidler et al., 1988, J. Immunol. 141:4053-4060.

Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Such antibodies can be produced using transgenic mice which are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. The transgenic mice are

immunized in the normal fashion with a selected antigen, e.g., all or a portion of a BPI of the invention. Monoclonal antibodies directed against the antigen can be obtained using conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA, IgM and IgE antibodies. For an overview of this technology for producing human antibodies, see Lonberg and Huszar (1995, Int. Rev. Immunol. 13:65-93). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, see, e.g., U.S. Patent 5,625,126; U.S. Patent 5,633,425; U.S. Patent 5,569,825; U.S. Patent 5,661,016; and U.S. Patent 5,545,806. In addition, companies such as Abgenix, Inc. (Freemont, CA) and Genpharm (San Jose, CA) can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, e.g., a mouse antibody, is used to guide the selection of a completely human antibody recognizing the same epitope. (Jespers et al. (1994) Bio/technology 12:899-903).

The antibodies of the present invention can also be generated using various phage display methods known in the art. In phage display methods, functional antibody domains are displayed on the surface of phage particles which carry the polynucleotide sequences encoding them. In a particular, such phage can be utilized to display antigen binding domains expressed from a repertoire or combinatorial antibody library (e.g., human or murine). Phage expressing an antigen binding domain that binds the antigen of interest can be selected or identified with antigen, e.g., using labeled antigen or antigen bound or captured to a solid surface or bead. Phage used in these methods are typically filamentous phage including fd and M13 binding domains expressed from phage with Fab, Fv or disulfide stabilized Fv antibody domains recombinantly fused to either the phage gene III or gene VIII protein. Phage display methods that can be used to make the antibodies of the present invention include those disclosed in Brinkman et al., J. Immunol. Methods 182:41-50 (1995); Ames et al., J. Immunol. Methods 184:177-186 (1995); Kettleborough et al., Eur. J. Immunol. 24:952-958 (1994); Persic et al., Gene 187

9-18 (1997); Burton et al., Advances in Immunology 57:191-280 (1994); PCT Application No. PCT/GB91/01134; PCT Publications WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; and U.S. Patent Nos. 5,698,426; 5,223,409; 5,403,484; 5,580,717; 5,427,908; 5,750,753; 5,821,047; 5,571,698; 5,427,908; 5,516,637; 5,780,225; 5,658,727; 5,733,743 and 5,969,108; each of which is incorporated herein by reference in its entirety.

As described in the above references, after phage selection, the antibody coding regions from the phage can be isolated and used to generate whole antibodies, including human antibodies, or any other desired antigen binding fragment, and expressed in any desired host, including mammalian cells, insect cells, plant cells, yeast, and bacteria, e.g., as described in detail below. For example, techniques to recombinantly produce Fab, Fab' and F(ab')2 fragments can also be employed using methods known in the art such as those disclosed in PCT publication WO 92/22324; Mullinax et al., BioTechniques 12(6):864-869 (1992); and Sawai et al., AJRI 34:26-34 (1995); and Better et al., Science 240:1041-1043 (1988) (said references incorporated by reference in their entireties).

Examples of techniques which can be used to produce single-chain Fvs and antibodies include those described in U.S. Patents 4,946,778 and 5,258,498; Huston et al., Methods in Enzymology 203:46-88 (1991); Shu et al., PNAS 90:7995-7999 (1993); and Skerra et al., Science 240:1038-1040 (1988).

The invention further provides for the use of bispecific antibodies, which can be made by methods known in the art. Traditional production of full length bispecific antibodies is based on the coexpression of two immunoglobulin heavy chain-light chain pairs, where the two chains have different specificities (Milstein et al., 1983, Nature 305:537-539). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of 10 different antibody molecules, of which only one has the correct bispecific structure. Purification of the correct molecule, which is usually done by affinity chromatography steps, is rather cumbersome, and the product yields are low. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., 1991, EMBO J. 10:3655-3659 .

According to a different and more preferred approach, antibody variable domains with the desired binding specificities (antibody-antigen combining sites) are fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy chain constant domain, comprising at least part of the hinge,

CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light chain binding, present in at least one of the fusions. DNAs encoding the immunoglobulin heavy chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. This provides for great flexibility in adjusting the mutual proportions of the three polypeptide fragments in embodiments when unequal ratios of the three polypeptide chains used in the construction provide the optimum yields. It is, however, possible to insert the coding sequences for two or all three polypeptide chains in one expression vector when the expression of at least two polypeptide chains in equal ratios results in high yields or when the ratios are of no particular significance.

In a preferred embodiment of this approach, the bispecific antibodies are composed of a hybrid immunoglobulin heavy chain with a first binding specificity in one arm, and a hybrid immunoglobulin heavy chain-light chain pair (providing a second binding specificity) in the other arm. It was found that this asymmetric structure facilitates the separation of the desired bispecific compound from unwanted immunoglobulin chain combinations, as the presence of an immunoglobulin light chain in only one half of the bispecific molecule provides for a facile way of separation. This approach is disclosed in WO 94/04690 published March 3,1994. For further details for generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology,1986, 121:210.

The invention provides functionally active fragments, derivatives or analogs of the anti-BPI immunoglobulin molecules. Functionally active means that the fragment, derivative or analog is able to elicit anti-anti-idiotype antibodies (i.e., tertiary antibodies) that recognize the same antigen that is recognized by the antibody from which the fragment, derivative or analog is derived. Specifically, in a preferred embodiment the antigenicity of the idiotype of the immunoglobulin molecule may be enhanced by deletion of framework and CDR sequences that are C-terminal to the CDR sequence that specifically recognizes the antigen. To determine which CDR sequences bind the antigen, synthetic peptides containing the CDR sequences can be used in binding assays with the antigen by any binding assay method known in the art.

The present invention provides antibody fragments such as, but not limited to, F(ab')2 fragments and Fab fragments. Antibody fragments which recognize specific epitopes may be generated by known techniques. F(ab')2 fragments consist of the variable region,

the light chain constant region and the CH1 domain of the heavy chain and are generated by pepsin digestion of the antibody molecule. Fab fragments are generated by reducing the disulfide bridges of the F(ab')2 fragments. The invention also provides heavy chain and light chain dimers of the antibodies of the invention, or any minimal fragment thereof such as Fvs or single chain antibodies (SCAs) (e.g., as described in U.S. Patent 4,946,778; Bird, 1988, Science 242:423-42; Huston et al., 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; and Ward et al., 1989, Nature 334:544-54), or any other molecule with the same specificity as the antibody of the invention. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide. Techniques for the assembly of functional Fv fragments in E. coli may be used (Skerra et al., 1988, Science 242:1038-1041).

In other embodiments, the invention provides fusion proteins of the immunoglobulins of the invention (or functionally active fragments thereof), for example in which the immunoglobulin is fused via a covalent bond (e.g., a peptide bond), at either the N-terminus or the C-terminus to an amino acid sequence of another protein (or portion thereof, preferably at least 10, 20 or 50 amino acid portion of the protein) that is not the immunoglobulin. Preferably the immunoglobulin, or fragment thereof, is covalently linked to the other protein at the N-terminus of the constant domain. As stated above, such fusion proteins may facilitate purification, increase half-life in vivo, and enhance the delivery of an antigen across an epithelial barrier to the immune system.

The immunoglobulins of the invention include analogs and derivatives that are either modified, i.e, by the covalent attachment of any type of molecule as long as such covalent attachment that does not impair immunospecific binding. For example, but not by way of limitation, the derivatives and analogs of the immunoglobulins include those that have been further modified, e.g., by glycosylation, acetylation, pegylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other protein, etc. Any of numerous chemical modifications may be carried out by known techniques, including, but not limited to specific chemical cleavage, acetylation, formylation, etc. Additionally, the analog or derivative may contain one or more non-classical amino acids.

The foregoing antibodies can be used in methods known in the art relating to the localization and activity of the BPIs of the invention, e.g., for imaging these proteins,

measuring levels thereof in appropriate physiological samples, in diagnostic methods, for radioimaging diagnosis or radioimmunotherapy, etc.

4.8. Expression Of Antibodies

The antibodies of the invention can be produced by any method known in the art for the synthesis of antibodies, in particular, by chemical synthesis or by recombinant expression, and are preferably produced by recombinant expression technique.

Recombinant expression of antibodies, or fragments, derivatives or analogs thereof, requires construction of a nucleic acid that encodes the antibody. If the nucleotide sequence of the antibody is known, a nucleic acid encoding the antibody may be assembled from chemically synthesized oligonucleotides (e.g., as described in Kutmeier et al., 1994, BioTechniques 17:242), which, briefly, involves the synthesis of overlapping oligonucleotides containing portions of the sequence encoding antibody, annealing and ligation of those oligonucleotides, and then amplification of the ligated oligonucleotides by PCR.

Alternatively, the nucleic acid ~~encoding the antibody may be obtained by cloning the~~ antibody. If a clone containing the nucleic acid encoding the particular antibody is not available, but the sequence of the antibody molecule is known, a nucleic acid encoding the antibody may be obtained from a suitable source (e.g., an antibody cDNA library, or cDNA library generated from any tissue or cells expressing the antibody) by PCR amplification using synthetic primers hybridizable to the 3' and 5' ends of the sequence or by cloning using an oligonucleotide probe specific for the particular gene sequence.

If an antibody molecule that specifically recognizes a particular antigen is not available (or a source for a cDNA library for cloning a nucleic acid encoding such an antibody), antibodies specific for a particular antigen may be generated by any method known in the art, for example, by immunizing an animal, such as a rabbit, to generate polyclonal antibodies or, more preferably, by generating monoclonal antibodies. Alternatively, a clone encoding at least the Fab portion of the antibody may be obtained by screening Fab expression libraries (e.g., as described in Huse et al., 1989, Science 246:1275-1281) for clones of Fab fragments that bind the specific antigen or by screening antibody libraries (See, e.g., Clackson et al., 1991, Nature 352:624; Hane et al., 1997 Proc. Natl. Acad. Sci. USA 94:4937).

Once a nucleic acid encoding at least the variable domain of the antibody molecule is obtained, it may be introduced into a vector containing the nucleotide sequence encoding the constant region of the antibody molecule (see, e.g., PCT Publication WO 86/05807; PCT Publication WO 89/01036; and U.S. Patent No. 5,122,464). Vectors containing the complete light or heavy chain for co-expression with the nucleic acid to allow the expression of a complete antibody molecule are also available. Then, the nucleic acid encoding the antibody can be used to introduce the nucleotide substitution(s) or deletion(s) necessary to substitute (or delete) the one or more variable region cysteine residues participating in an intrachain disulfide bond with an amino acid residue that does not contain a sulfhydyl group. Such modifications can be carried out by any method known in the art for the introduction of specific mutations or deletions in a nucleotide sequence, for example, but not limited to, chemical mutagenesis, in vitro site directed mutagenesis (Hutchinson et al., 1978, J. Biol. Chem. 253:6551), PCT based methods, etc. In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., 1984, Proc. Natl. Acad. Sci. 81:851-855; Neuberger et al., 1984, Nature 312:604-608; Takeda et al., 1985, Nature 314:452-454) by splicing genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. As described supra, a chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human antibody constant region, e.g., humanized antibodies

Once a nucleic acid encoding an antibody molecule of the invention has been obtained, the vector for the production of the antibody molecule may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing the protein of the invention by expressing nucleic acid containing the antibody molecule sequences are described herein. Methods which are well known to those skilled in the art can be used to construct expression vectors containing an antibody molecule coding sequences and appropriate transcriptional and translational control signals. These methods include, for example, in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. See, for example, the techniques described in Sambrook et al. (1990, Molecular Cloning, A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) and Ausubel et al. (eds., 1998, Current Protocols in Molecular Biology, John Wiley & Sons, NY).

The expression vector is transferred to a host cell by conventional techniques and the transfected cells are then cultured by conventional techniques to produce an antibody of the invention.

The host cells used to express a recombinant antibody of the invention may be either bacterial cells such as Escherichia coli, or, preferably, eukaryotic cells, especially for the expression of whole recombinant antibody molecule. In particular, mammalian cells such as Chinese hamster ovary cells (CHO), in conjunction with a vector such as the major intermediate early gene promoter element from human cytomegalovirus is an effective expression system for antibodies (Foecking et al., 198, Gene 45:101; Cockett et al., 1990, Bio/Technology 8:2).

A variety of host-expression vector systems may be utilized to express an antibody molecule of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, express the antibody molecule of the invention in situ. These include but are not limited to microorganisms such as bacteria (e.g., E. coli, B. subtilis) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing antibody coding sequences; yeast (e.g., Saccharomyces, Pichia) transformed with recombinant yeast expression vectors containing antibody coding sequences; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing the antibody coding sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing antibody coding sequences; or mammalian cell systems (e.g., COS, CHO, BHK, 293, 3T3 cells) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the antibody molecule being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions comprising an antibody molecule, vectors which direct the expression of high levels of fusion protein products that are readily purified may be

desirable. Such vectors include, but are not limited, to the E. coli expression vector pUR278 (Ruther et al., 1983, EMBO J. 2:1791), in which the antibody coding sequence may be ligated individually into the vector in frame with the lac Z coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, 1985, Nucleic Acids Res. 13:3101-3109; Van Heeke & Schuster, 1989, J. Biol. Chem. 24:5503-5509); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption and binding to a matrix glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

In an insect system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in Spodoptera frugiperda cells. The antibody coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). In mammalian host cells, a number of viral-based expression systems (e.g., an adenovirus expression system) may be utilized. As discussed above, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein.

For long-term, high-yield production of recombinant antibodies, stable expression is preferred. For example, cells lines that stably express an antibody of interest can be produced by transfecting the cells with an expression vector comprising the nucleotide sequence of the antibody and the nucleotide sequence of a selectable (e.g., neomycin or hygromycin), and selecting for expression of the selectable marker. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that interact directly or indirectly with the antibody molecule.

The expression levels of the antibody molecule can be increased by vector amplification (for a review, see Bebbington and Hentschel, The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells in DNA cloning, Vol.3. (Academic Press, New York, 1987)). When a marker in the vector system expressing antibody is amplifiable, increase in the level of inhibitor present in culture of

host cell will increase the number of copies of the marker gene. Since the amplified region is associated with the antibody gene, production of the antibody will also increase (Crouse et al., 1983, Mol. Cell. Biol. 3:257).

The host cell may be co-transfected with two expression vectors of the invention, the first vector encoding a heavy chain derived polypeptide and the second vector encoding a light chain derived polypeptide. The two vectors may contain identical selectable markers which enable equal expression of heavy and light chain polypeptides. Alternatively, a single vector may be used which encodes both heavy and light chain polypeptides. In such situations, the light chain should be placed before the heavy chain to avoid an excess of toxic free heavy chain (Proudfoot, 1986, Nature 322:52; Kohler, 1980, Proc. Natl. Acad. Sci. USA 77:2197). The coding sequences for the heavy and light chains may comprise cDNA or genomic DNA.

Once the antibody molecule of the invention has been recombinantly expressed, it may be purified by any method known in the art for purification of an antibody molecule, for example, by chromatography (e.g., ion exchange chromatography, affinity chromatography such as with protein A or specific antigen, and sizing column chromatography), centrifugation, differential solubility, or by any other standard technique for the purification of proteins.

Alternatively, any fusion protein may be readily purified by utilizing an antibody specific for the fusion protein being expressed. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht et al., 1991, Proc. Natl. Acad. Sci. USA 88:8972-897). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the open reading frame of the gene is translationally fused to an amino-terminal tag consisting of six histidine residues. The tag serves as a matrix binding domain for the fusion protein. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni2+ nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

4.9. Conjugated Antibodies

In a preferred embodiment, anti-BPI antibodies or fragments thereof are conjugated to a diagnostic or therapeutic moiety. The antibodies can be used for diagnosis or to assess the effect of tumour growth or metastatic spread or to determine the efficacy of a given

treatment regimen. Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, radioactive nuclides, positron emitting metals (for use in positron emission tomography), and nonradioactive paramagnetic metal ions. See generally U.S. Patent No. 4,741,900 for metal ions which can be conjugated to antibodies for use as diagnostics according to the present invention. Suitable enzymes include horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; suitable prosthetic groups include streptavidin, avidin and biotin; suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride and phycoerythrin; suitable luminescent materials include luminol; suitable bioluminescent materials include luciferase, luciferin, and aequorin; and suitable radioactive nuclides include 125I, 131I, 111In and 99Tc.

An anti-BPI antibodies or fragments thereof can be conjugated to a therapeutic agent or drug moiety to modify a given biological response. The therapeutic agent or drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, gamma-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator, a thrombotic agent or an anti-angiogenic agent, e.g., angiostatin or endostatin; or, a biological response modifier such as a lymphokine, interleukin-1 (IL-1), interleukin-2 (IL-2), interleukin-6 (IL-6), granulocyte macrophage colony stimulating factor (GM-breast tissue), granulocyte colony stimulating factor (G-breast tissue), nerve growth factor (NGF) or other growth factor.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, e.g., Arnon et al., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld et al. (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom et al., "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson et al. (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera et al. (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The

Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin et al. (eds.), pp. 303-16 (Academic Press 1985), and Thorpe et al., "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", Immunol. Rev., 62:119-58 (1982).
Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.
An antibody with or without a therapeutic moiety conjugated to it can be used as a therapeutic that is administered alone or in combination with cytotoxic factor(s) and/or cytokine(s).

4.10. Diagnosis of breast cancer

In accordance with the invention, test samples of tissue or body fluids (e.g., serum, plasma or urine) obtained from a subject suspected of having or known to have breast cancer can be used for diagnosis or monitoring. In one embodiment, a decreased abundance of one or more BFs or BPIs (or any combination of them) in a test sample relative to a control sample (from a subject or subjects free from breast cancer) or a previously determined reference range indicates the presence of breast cancer; BFs and BPIs suitable for this purpose are identified in Tables I and III, respectively. In another embodiment of the invention, an increased abundance of one or more BFs or BPIs (or any combination of them) in a test sample compared to a control sample or a previously determined reference range indicates the presence of breast cancer; BFs and BPIs suitable for this purpose are identified in Tables II and IV, respectively. In another embodiment, the relative abundance of one or more BFs or BPIs (or any combination of them) in a test sample compared to a control sample or a previously determined reference range indicates a subtype of breast cancer (e.g., familial or sporadic breast cancer). In yet another embodiment, the relative abundance of one or more BFs or BPIs (or any combination of them) in a test sample relative to a control sample or a previously determined reference range indicates the degree or severity of breast cancer. In any of the aforesaid methods, detection of one or more BPIs described herein may optionally be combined with detection of one or more additional biomarkers for breast cancer. Many methods standard in the art can be employed to measure the level of BFs and BPIs, including but not limited to the Preferred Technology described herein, kinase assays, immunoassays to detect and/or visualize the BPI (e.g., Western blot, immunoprecipitation

followed by sodium dodecyl sulfate polyacrylamide gel electrophoresis, immunocytochemistry, etc.). In cases where a BPI has a known function, an assay for that function may be used to measure BPI expression. In a further embodiment, a decreased abundance of mRNA including one or more BPIs identified in Table III (or any combination of them) in a test sample relative to a control sample or a previously determined reference range indicates the presence of breast cancer. In yet a further embodiment, an increased abundance of mRNA encoding one or more BPIs identified in Table IV (or any combination of them) in a test sample relative to a control sample or previously determined reference range indicates the presence of breast cancer. Hybridization assays can be used to detect BPI expression by detecting and/or visualizing mRNA encoding the BPI (e.g., Northern assays, dot blots, in situ hybridization, etc.). In another embodiment of the invention, labeled antibodies, derivatives and analogs thereof, which specifically bind to a BPI can be used for diagnostic purposes to detect, diagnose, or monitor breast cancer. Preferably, breast cancer is detected in a mammal and most preferably in a human. Such antibodies can be of use in vitro or in vivo as, for example, in radioimaging diagnosis.

4.11. Therapeutic Use Of BPIs

The invention provides for treatment or prevention of metastatic breast cancer by administration of a therapeutic compound. Such compounds include but are not limited to: BPIs and analogs and derivatives (including fragments) thereof; antibodies thereto; nucleic acids encoding BPIs, analogs, or derivatives; antisense nucleic acids to a gene encoding a BPI, and agonists and antagonists of a gene encoding a BPI. An important feature of the present invention is the identification of genes encoding BPIs involved in breast cancer. Breast cancer can be treated or prevented by local administration of a therapeutic compound that promotes function or expression of one or more BPIs that are decreased in the tissue of breast cancer patients, or by administration of a therapeutic compound that reduces function or expression of one or more BPIs that are increased in the tissue of breast cancer patients.

In one embodiment, one or more antibodies each specifically binding to a BPI are administered alone or in combination with one or more additional therapeutic compounds or treatments. Examples of such therapeutic compounds or treatments include, but are

not limited to, taxol, cyclophosphamide, tamoxifen, fluorouracil, methotrexate and doxorubicin.

Preferably, a biological product such as an antibody is allogeneic to the subject to which it is administered. In a preferred embodiment, a human BPI or derivative thereof, a nucleotide sequence encoding a human BPI, or an antibody to a human BPI, is administered to a human patient for therapy or prophylaxis.

Preferably, a biological product such as an antibody is delivered to a subject by localised administration such that the biological product reacts with breast cancer cells which have spread to sites distal to normal breast tissue.

4.11.1. Treatment And Prevention Of Breast Cancer

Breast cancer is treated or prevented by administration to a subject suspected of having or known to have breast cancer or to be at risk of developing breast cancer of a compound that modulates (i.e., increases or decreases) the level or activity (i.e., function) of one or more BPIs - or the level of one or more BFs - that are differentially present in the tissue of subjects having breast cancer compared with tissue of subjects free from breast cancer. In one embodiment, breast cancer is treated or prevented by administering to a subject suspected of having or known to have breast cancer or to be at risk of developing breast cancer a compound that upregulates (i.e., increases) the level or activity (i.e., function) of one or more BPIs - or the level of one or more BFs - that are decreased in the breast tissue of subjects having breast cancer. In another embodiment, a compound is administered that upregulates the level or activity (i.e., function) of one or more BPIs - or the level of one or more BFs - that are increased in the tissue of subjects having breast cancer. Examples of such a compound include but are not limited to: BPIs, derivatives or fragments thereof that are functionally active (e.g., in in vitro assays or in animal models as described above), nucleic acids encoding a BPI or functionally active derivative or fragment thereof (e.g., for use in gene therapy), and, for those BPIs with enzymatic activity, compounds or molecules known to modulate that enzymatic activity. Other compounds that can be used, e.g., BPI agonists, can be identified using in vitro assays. Breast cancer is also treated or prevented by administration to a subject suspected of having or known to have breast cancer or to be at risk of developing breast cancer of a compound that downregulates the level or activity of one or more BPIs - or the level of one or more BFs - that are increased in the tissue of subjects having breast cancer. In

another embodiment, a compound is administered that downregulates the level or activity of one or more BPIs - or the level of one or more BFs - that are decreased in the tissue of subjects having breast cancer. Examples of such a compound include but are not limited to BPI anti-sense oligonucleotides, ribozymes, or antibodies directed against BPIs, as well as compounds that inhibit the enzymatic activity of a BPI , e.g. BPIs 34, 39, 40, 48, 51. Other compounds that can be used, e.g., BPI antagonists and small molecule BPI antagonists, can be identified using in vitro assays.

In a preferred embodiment, therapy or prophylaxis is tailored to the needs of an individual subject. Thus, in specific embodiments, compounds that promote the level or function of one or more BPIs, or the level of one or more BFs, are therapeutically or prophylactically administered to a subject suspected of having or known to have breast cancer, in whom the levels or functions of said one or more BPIs, or levels of said one or more BFs, are absent or are decreased relative to a control or normal reference range. In further embodiments, compounds that promote the level or function of one or more BPIs, or the level of one or more BFs, are therapeutically or prophylactically administered to a subject suspected of having or known to have breast cancer in whom the levels or functions of said one or more BPIs, or levels of said one or more BFs, are increased relative to a control or to a reference range. In further embodiments, compounds that decrease the level or function of one or more BPIs, or the level of one or more BFs, are therapeutically or prophylactically administered to a subject suspected of having or known to have breast cancer in whom the levels or functions of said one or more BPIs, or levels of said one or more BFs, are increased relative to a control or to a reference range. In further embodiments, compounds that decrease the level or function of one or more BPIs, or the level of one or more BFs, are therapeutically or prophylactically administered to a subject suspected of having or known to have breast cancer in whom the levels or functions of said one or more BPIs, or levels of said one or more BFs, are decreased relative to a control or to a reference range. The change in BPI function or level, or BF level, due to the administration of such compounds can be readily detected, e.g., by obtaining a tissue sample (e.g., from biopsy tissue) and assaying in vitro the levels of said BFs or the levels or activities of said BPIs, or the levels of mRNAs encoding said BPIs. or any combination of the foregoing. Such assays can be performed before and after the administration of the compound as described herein.

In a referred embodiment, a compound which modualtes the level or function of one or more BPIs is administered in a localised fashion such that the compound binds or reacts with cancer cells which have spread to a site distal to normal breast tissue, such as cancer cells which have spread to the circulatory system or to the liver or to bone tissue.
The compounds of the invention include but are not limited to any compound, e.g., a small organic molecule, protein, peptide, antibody, nucleic acid, etc. that restores the cancer tissue BPI or BF profile towards normal with the proviso that such compound is not already the subject of a patent claim for restoration of a particular metastatic breast cancer tissue BPI or BF profile described herein.

4.11.2. Gene Therapy

In a specific embodiment, nucleic acids comprising a sequence encoding a BPI or functional derivative thereof, are administered to a patient in a localised fashion to modulate BPI function by way of gene therapy. Gene therapy refers to administration to a subject of an expressed or expressible nucleic acid. In this embodiment, the nucleic acid produces its encoded protein that mediates a therapeutic effect by modulating BPI function.
Any of the methods for gene therapy available in the art can be used according to the present invention. Exemplary methods are described below.
For general reviews of the methods of gene therapy, see Goldspiel et al., 1993, Clinical Pharmacy 12:488-505; Wu and Wu, 1991, Biotherapy 3:87-95; Tolstoshev, 1993, Ann. Rev. Pharmacol. Toxicol. 32:573-596; Mulligan, 1993, Science 260:926-932; and Morgan and Anderson, 1993, Ann. Rev. Biochem. 62:191-217; May, 1993, TIBTECH 11(5):155-215). Methods commonly known in the art of recombinant DNA technology which can be used are described in Ausubel et al. (eds.), 1993, Current Protocols in Molecular Biology, John Wiley & Sons, NY; and Kriegler, 1990, Gene Transfer and Expression, A Laboratory Manual, Stockton Press, NY.
In a preferred aspect, the compound comprises a nucleic acid encoding a BPI or fragment or chimeric protein thereof, said nucleic acid being part of an expression vector that expresses a BPI or fragment or chimeric protein thereof in a suitable host. In particular, such a nucleic acid has a promoter operably linked to the BPI coding region, said promoter being inducible or constitutive (and, optionally, tissue-specific). In another particular embodiment, a nucleic acid molecule is used in which the BPI coding

sequences and any other desired sequences are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intrachromosomal expression of the BPI nucleic acid (Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature 342:435-438). Delivery of the nucleic acid into a patient may be direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid-carrying vector; this approach is known as in vivo gene therapy. Alternatively, delivery of the nucleic acid into the patient may be indirect, in which case cells are first transformed with the nucleic acid in vitro and then transplanted into the patient; this approach is known as ex vivo gene therapy. In a specific embodiment, the nucleic acid is directly administered in vivo, where it is expressed to produce the encoded product. This can be accomplished by any of numerous methods known in the art, e.g., by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, e.g., by infection using a defective or attenuated retroviral or other viral vector (see U.S. Patent No. 4,980,286); by direct injection of naked DNA; by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont); by coating with lipids, cell-surface receptors or transfecting agents; by encapsulation in liposomes, microparticles or microcapsules; by administering it in linkage to a peptide which is known to enter the nucleus; or by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see, e.g., Wu and Wu, 1987, J. Biol. Chem. 262:4429-4432), which can be used to target cell types specifically expressing the receptors. In another embodiment, a nucleic acid-ligand complex can be formed in which the ligand comprises a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted in vivo for cell specific uptake and expression, by targeting a specific receptor (see, e.g., PCT Publications WO 92/06180 dated April 16, 1992 (Wu et al.); WO 92/22635 dated December 23, 1992 (Wilson et al.); WO92/20316 dated November 26, 1992 (Findeis et al.); WO93/14188 dated July 22, 1993 (Clarke et al.), WO 93/20221 dated October 14, 1993 (Young)). Alternatively, the nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination (Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature 342:435-438).

In a specific embodiment, a viral vector that contains a nucleic acid encoding a BPI is used. For example, a retroviral vector can be used (see Miller et al., 1993, Meth. Enzymol. 217:581-599). These retroviral vectors have been modified to delete retroviral sequences that are not necessary for packaging of the viral genome and integration into host cell DNA. The nucleic acid encoding the BPI to be used in gene therapy is cloned into the vector, which facilitates delivery of the gene into a patient. More detail about retroviral vectors can be found in Boesen et al., 1994, Biotherapy 6:291-302, which describes the use of a retroviral vector to deliver the mdr1 gene to hematopoietic stem cells in order to make the stem cells more resistant to chemotherapy. Other references illustrating the use of retroviral vectors in gene therapy are: Clowes et al., 1994, J. Clin. Invest. 93:644-651; Kiem et al., 1994, Blood 83:1467-1473; Salmons and Gunzberg, 1993, Human Gene Therapy 4:129-141; and Grossman and Wilson, 1993, Curr. Opin. in Genetics and Devel. 3:110-114.

Adenoviruses are other viral vectors that can be used in gene therapy. Adenoviruses are especially attractive vehicles for delivering genes to respiratory epithelia. Adenoviruses naturally infect respiratory epithelia where they cause a mild disease. Other targets for adenovirus-based delivery systems are liver, the central nervous system, endothelial cells, and muscle. Adenoviruses have the advantage of being capable of infecting non-dividing cells. Kozarsky and Wilson, 1993, Current Opinion in Genetics and Development 3:499-503 present a review of adenovirus-based gene therapy. Bout et al., 1994, Human Gene Therapy 5:3-10 demonstrated the use of adenovirus vectors to transfer genes to the respiratory epithelia of rhesus monkeys. Other instances of the use of adenoviruses in gene therapy can be found in Rosenfeld et al., 1991, Science 252:431-434; Rosenfeld et al., 1992, Cell 68:143-155; Mastrangeli et al., 1993, J. Clin. Invest. 91:225-234; PCT Publication WO94/12649; and Wang, et al., 1995, Gene Therapy 2:775-783.

Adeno-associated virus (AAV) has also been proposed for use in gene therapy (Walsh et al., 1993, Proc. Soc. Exp. Biol. Med. 204:289-300; U.S. Patent No. 5,436,146).

Another approach to gene therapy involves transferring a gene to cells in tissue culture by such methods as electroporation, lipofection, calcium phosphate mediated transfection, or viral infection. Usually, the method of transfer includes the transfer of a selectable marker to the cells. The cells are then placed under selection to isolate those cells that have taken up and are expressing the transferred gene. Those cells are then delivered to a patient.

In this embodiment, the nucleic acid is introduced into a cell prior to administration in vivo of the resulting recombinant cell. Such introduction can be carried out by any method known in the art, including but not limited to transfection, electroporation, microinjection, infection with a viral or bacteriophage vector containing the nucleic acid sequences, cell fusion, chromosome-mediated gene transfer, microcell-mediated gene transfer, spheroplast fusion, etc. Numerous techniques are known in the art for the introduction of foreign genes into cells (see, e.g., Loeffler and Behr, 1993, Meth. Enzymol. 217:599-618; Cohen et al., 1993, Meth. Enzymol. 217:618-644; Cline, 1985, Pharmac. Ther. 29:69-92) and may be used in accordance with the present invention, provided that the necessary developmental and physiological functions of the recipient cells are not disrupted. The technique should provide for the stable transfer of the nucleic acid to the cell, so that the nucleic acid is expressible by the cell and preferably heritable and expressible by its cell progeny.

The resulting recombinant cells can be delivered to a patient by various methods known in the art. In a preferred embodiment, epithelial cells are injected, e.g., subcutaneously. In another embodiment, recombinant skin cells may be applied as a skin graft onto the patient. Recombinant blood cells (e.g., hematopoietic stem or progenitor cells) are preferably administered intravenously. The amount of cells envisioned for use depends on the desired effect, patient state, etc., and can be determined by one skilled in the art. Cells into which a nucleic acid can be introduced for purposes of gene therapy encompass any desired, available cell type, and include but are not limited to neuronal cells, glial cells (e.g., oligodendrocytes or astrocytes), epithelial cells, endothelial cells, keratinocytes, fibroblasts, muscle cells, hepatocytes; blood cells such as T lymphocytes, B lymphocytes, monocytes, macrophages, neutrophils, eosinophils, megakaryocytes, granulocytes; various stem or progenitor cells, in particular hematopoietic stem or progenitor cells, e.g., as obtained from bone marrow, umbilical cord blood, peripheral blood or fetal liver.

In a preferred embodiment, the cell used for gene therapy is autologous to the patient.

In an embodiment in which recombinant cells are used in gene therapy, a nucleic acid encoding a BPI is introduced into the cells such that it is expressible by the cells or their progeny, and the recombinant cells are then administered in vivo for therapeutic effect. In a specific embodiment, stem or progenitor cells are used. Any stem or progenitor cells which can be isolated and maintained in vitro can be used in accordance with this

embodiment of the present invention (see e.g. PCT Publication WO 94/08598, dated April 28, 1994; Stemple and Anderson, 1992, Cell 71:973-985; Rheinwald, 1980, Meth. Cell Bio. 21A:229; and Pittelkow and Scott, 1986, Mayo Clinic Proc. 61:771).

In a specific embodiment, the nucleic acid to be introduced for purposes of gene therapy comprises an inducible promoter operably linked to the coding region, such that expression of the nucleic acid is controllable by controlling the presence or absence of the appropriate inducer of transcription.

Direct injection of a DNA coding for a BPI may also be performed according to, for example, the techniques described in United States Patent No. 5,589,466. These techniques involve the injection of "naked DNA", i.e., isolated DNA molecules in the absence of liposomes, cells, or any other material besides a suitable carrier. The injection of DNA encoding a protein and operably linked to a suitable promoter results in the production of the protein in cells near the site of injection and the elicitation of an immune response in the subject to the protein encoded by the injected DNA. In a preferred embodiment, naked DNA comprising (a) DNA encoding a BPI and (b) a promoter are injected into a subject to elicit an immune response to the BPI.

4.11.3 Antisense Regulation of BPIs

In a specific embodiment, BPI expression is inhibited by use of BPI antisense nucleic acids. The present invention provides the therapeutic or prophylactic use of nucleic acids comprising at least six nucleotides that are antisense to a gene or cDNA encoding a BPI or a portion thereof. As used herein, a BPI "antisense" nucleic acid refers to a nucleic acid capable of hybridizing by virtue of some sequence complementarity to a portion of an RNA (preferably mRNA) encoding a BPI. The antisense nucleic acid may be complementary to a coding and/or noncoding region of an mRNA encoding a BPI. Such antisense nucleic acids have utility as compounds that inhibit BPI expression, and can be used in the treatment or prevention of breast cancer.

The antisense nucleic acids of the invention are double-stranded or single-stranded oligonucleotides, RNA or DNA or a modification or derivative thereof, and can be directly administered to a cell or produced intracellularly by transcription of exogenous, introduced sequences.

The invention further provides pharmaceutical compositions comprising an effective amount of the BPI antisense nucleic acids of the invention in a pharmaceutically acceptable carrier, as described infra.

In another embodiment, the invention provides methods for inhibiting the expression of a BPI nucleic acid sequence in a prokaryotic or eukaryotic cell comprising providing the cell with an effective amount of a composition comprising a BPI antisense nucleic acid of the invention.

BPI antisense nucleic acids and their uses are described in detail below.

4.11.3.1 BPI Antisense Nucleic Acids

The BPI antisense nucleic acids are of at least six nucleotides and are preferably oligonucleotides ranging from 6 to about 50 oligonucleotides. In specific aspects, the oligonucleotide is at least 10 nucleotides, at least 15 nucleotides, at least 100 nucleotides, or at least 200 nucleotides. The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof and can be single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone. The oligonucleotide may include other appended groups such as peptides; agents that facilitate transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. USA 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. WO 88/09810, published December 15, 1988); hybridization-triggered cleavage agents (see, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents (see, e.g., Zon, 1988, Pharm. Res. 5:539-549).

In a preferred aspect of the invention, a BPI antisense oligonucleotide is provided, preferably of single-stranded DNA. The oligonucleotide may be modified at any position on its structure with substituents generally known in the art.

The BPI antisense oligonucleotide may comprise at least one of the following modified base moieties: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-

methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5?-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, 2,6-diaminopurine, and other base analogs.

In another embodiment, the oligonucleotide comprises at least one modified sugar moiety, e.g., one of the following sugar moieties: arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the oligonucleotide comprises at least one of the following modified phosphate backbones: a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, a formacetal, or an analog of formacetal.

In yet another embodiment, the oligonucleotide is an ?-anomeric oligonucleotide. An ?-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual ?-units, the strands run parallel to each other (Gautier et al., 1987, Nucl. Acids Res. 15:6625-6641).

The oligonucleotide may be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent. Oligonucleotides of the invention may be synthesized by standard methods known in the art, e.g., by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein et al. (1988, Nucl. Acids Res. 16:3209), and methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., 1988, Proc. Natl. Acad. Sci. USA 85:7448-7451).

In a specific embodiment, the BPI antisense nucleic acid of the invention is produced intracellularly by transcription from an exogenous sequence. For example, a vector can be introduced in vivo such that it is taken up by a cell, within which cell the vector or a portion thereof is transcribed, producing an antisense nucleic acid (RNA) of the invention. Such a vector would contain a sequence encoding the BPI antisense nucleic acid. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be

constructed by recombinant DNA technology standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the BPI antisense RNA can be by any promoter known in the art to act in mammalian, preferably human, cells. Such promoters can be inducible or constitutive. Examples of such promoters are outlined above.
The antisense nucleic acids of the invention comprise a sequence complementary to at least a portion of an RNA transcript of a gene encoding a BPI, preferably a human gene encoding a BPI. However, absolute complementarity, although preferred, is not required. A sequence "complementary to at least a portion of an RNA," as referred to herein, means a sequence having sufficient complementarity to be able to hybridize under stringent conditions (e.g., highly stringent conditions comprising hybridization in 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65oC and washing in 0.1xSSC/0.1% SDS at 68oC, or moderately stringent conditions comprising washing in 0.2xSSC/0.1% SDS at 42oC) with the RNA, forming a stable duplex; in the case of double-stranded BPI antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA encoding a BPI it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

4.11.3.2 Therapeutic Use Of BPI Antisense Nucleic Acids

The BPI antisense nucleic acids can be used to treat or prevent breast cancer when the target BPI is overexpressed in the tissue of patients suspected of having or suffering from metastatic breast cancer. In a preferred embodiment, a single-stranded DNA antisense BPI oligonucleotide is used. In another preferred embodiment, a single-stranded DNA antisense BPI oligonucleotide is delivered to a subject by localised administration such that the single-stranded DNA antisense BPI oligonucleotide reacts with breast cancer cells which have spread to sites distal to normal breast tissue.

Cell types which express or overexpress RNA encoding a BPI can be identified by various methods known in the art. Such cell types include but are not limited to breast epithelial cells (luminal or myoepithelial cells). Such methods include, but are not limited to, hybridization with a BPI-specific nucleic acid (e.g., by Northern hybridization, dot blot hybridization, in situ hybridization), observing the ability of RNA from the cell type to be translated in vitro into a BPI, immunoassay, etc. In a preferred aspect, primary tissue from a patient can be assayed for BPI expression prior to treatment, e.g., by immunocytochemistry or in situ hybridization.

Pharmaceutical compositions of the invention, comprising an effective amount of a BPI antisense nucleic acid in a pharmaceutically acceptable carrier, can be administered to a patient having breast cancer.

The amount of BPI antisense nucleic acid which will be effective in the treatment of breast cancer can be determined by standard clinical techniques.

In a specific embodiment, pharmaceutical compositions comprising one or more BPI antisense nucleic acids are administered via liposomes, microparticles, or microcapsules. In various embodiments of the invention, such compositions may be used to achieve sustained release of the BPI antisense nucleic acids. In a specific embodiment, it may be desirable to use liposomes targeted via antibodies to specific identifiable tumor antigens (Leonetti et al., 1990, Proc. Natl. Acad. Sci. USA 87:2448-2451; Renneisen et al., 1990, J. Biol. Chem. 265:16337-16342).

4.11.3.3 Inhibitory Ribozyme And Triple Helix Approaches

In another embodiment, symptoms of metastatic breast cancer may be ameliorated by decreasing the level of a BPI or BPI activity by using gene sequences encoding the BPI in conjunction with well-known gene "knock-out," ribozyme or triple helix methods to decrease gene expression of a BPI. In this approach ribozyme or triple helix molecules are used to modulate the activity, expression or synthesis of the gene encoding the BPI, and thus to ameliorate the symptoms of breast cancer. Such molecules may be designed to reduce or inhibit expression of a mutant or non-mutant target gene. Techniques for the production and use of such molecules are well known to those of skill in the art.

Ribozyme molecules designed to catalytically cleave gene mRNA transcripts encoding a BPI can be used to prevent translation of target gene mRNA and, therefore, expression of

the gene product. (See, e.g., PCT International Publication WO90/11364, published October 4, 1990; Sarver et al., 1990, Science 247:1222-1225).

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. (For a review, see Rossi, 1994, Current Biology 4, 469-471). The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by an endonucleolytic cleavage event. The composition of ribozyme molecules must include one or more sequences complementary to the target gene mRNA, and must include the well known catalytic sequence responsible for mRNA cleavage. For this sequence, see, e.g., U.S. Patent No. 5,093,246, which is incorporated herein by reference in its entirety.

While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy mRNAs encoding a BPI, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Myers, 1995, Molecular Biology and Biotechnology: A Comprehensive Desk Reference, VCH Publishers, New York, (see especially Figure 4, page 833) and in Haseloff and Gerlach, 1988, Nature, 334, 585-591, each of which is incorporated herein by reference in its entirety.

Preferably the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the mRNA encoding the BPI, i.e., to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one that occurs naturally in Tetrahymena thermophila (known as the IVS, or L-19 IVS RNA) and that has been extensively described by Thomas Cech and collaborators (Zaug, et al., 1984, Science, 224, 574-578; Zaug and Cech, 1986, Science, 231, 470-475; Zaug, et al., 1986, Nature, 324, 429-433; published International patent application No. WO 88/04300 by University Patents Inc.; Been and Cech, 1986, Cell, 47, 207-216). The Cech-type ribozymes have an eight base pair active site which hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes which target eight base-pair active site sequences that are present in the gene encoding the BPI.

As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (e.g., for improved stability, targeting, etc.) and should be delivered to cells that express the BPI in vivo. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous mRNA encoding the BPI and inhibit translation. Because ribozymes, unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficacy.

Endogenous BPI expression can also be reduced by inactivating or "knocking out" the gene encoding the BPI, or the promoter of such a gene, using targeted homologous recombination (e.g., see Smithies, et al., 1985, Nature 317:230-234; Thomas and Capecchi, 1987, Cell 51:503-512; Thompson et al., 1989, Cell 5:313-321; and Zijlstra et al., 1989, Nature 342:435-438, each of which is incorporated by reference herein in its entirety). For example, a mutant gene encoding a non-functional BPI (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous gene (either the coding regions or regulatory regions of the gene encoding the BPI) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express the target gene in vivo. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the target gene. Such approaches are particularly suited in the agricultural field where modifications to ES (embryonic stem) cells can be used to generate animal offspring with an inactive target gene (e.g., see Thomas and Capecchi, 1987 and Thompson, 1989, supra). However this approach can be adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site in vivo using appropriate viral vectors.

Alternatively, the endogenous expression of a gene encoding a BPI can be reduced by targeting deoxyribonucleotide sequences complementary to the regulatory region of the gene (i.e., the gene promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene encoding the BPI in target cells in the body. (See generally, Helene, 1991, Anticancer Drug Des., 6(6), 569-584; Helene, et al., 1992, Ann. N.Y. Acad. Sci., 660, 27-36; and Maher, 1992, Bioassays 14(12), 807-815).

Nucleic acid molecules to be used in triplex helix formation for the inhibition of transcription should be single stranded and composed of deoxynucleotides. The base composition of these oligonucleotides must be designed to promote triple helix formation

via Hoogsteen base pairing rules, which generally require sizeable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may be pyrimidine-based, which will result in TAT and CGC+ triplets across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarity to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules may be chosen that are purine-rich, for example, contain a stretch of G residues. These molecules will form a triple helix with a DNA duplex that is rich in GC pairs, in which the majority of the purine residues are located on a single strand of the targeted duplex, resulting in GGC triplets across the three strands in the triplex.

Alternatively, the potential sequences that can be targeted for triple helix formation may be increased by creating a so called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

In instances wherein the antisense, ribozyme, or triple helix molecules described herein are utilized to inhibit mutant gene expression, it is possible that the technique may so efficiently reduce or inhibit the transcription (triple helix) or translation (antisense, ribozyme) of mRNA produced by normal gene alleles of a BPI that the situation may arise wherein the concentration of BPI present may be lower than is necessary for a normal phenotype. In such cases, to ensure that substantially normal levels of activity of a gene encoding a BPI are maintained, gene therapy may be used to introduce into cells nucleic acid molecules that encode and express the BPI that exhibit normal gene activity and that do not contain sequences susceptible to whatever antisense, ribozyme, or triple helix treatments are being utilized. Alternatively, in instances whereby the gene encodes an extracellular protein, normal BPI can be co-administered in order to maintain the requisite level of BPI activity.

Anti-sense RNA and DNA, ribozyme, and triple helix molecules of the invention may be prepared by any method known in the art for the synthesis of DNA and RNA molecules, as discussed above. These include techniques for chemically synthesizing oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding the antisense

RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors that incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

4.12 Therapeutic and Prophylactic Compositions and Their Use

The invention provides methods of treatment (and prophylaxis) comprising administering to a subject an effective amount of a compound of the invention. In a preferred aspect, the compound is substantially purified (e.g., substantially free from substances that limit its effect or produce undesired side-effects). The subject is preferably a mammal, including but not limited to animals such as cows, pigs, horses, chickens, cats, dogs, etc., and is most preferably human. In a specific embodiment, a non-human mammal is the subject.

Formulations and methods of administration that can be employed when the compound comprises a nucleic acid are described above; additional appropriate formulations and routes of administration are described below.

Various delivery systems are known and can be used to administer a compound of the invention, e.g., encapsulation in liposomes, microparticles, microcapsules, recombinant cells capable of expressing the compound, receptor-mediated endocytosis (see, e.g., Wu and Wu, 1987, J. Biol. Chem. 262:4429-4432), construction of a nucleic acid as part of a retroviral or other vector, etc. Methods of introduction can be enteral or parenteral and include but are not limited to intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, and oral routes. The compounds may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (e.g., oral mucosa, rectal and intestinal mucosa, etc.) and may be administered together with other biologically active agents. Administration can be systemic or local. Pulmonary administration can also be employed, e.g., by use of an inhaler or nebulizer, and formulation with an aerosolizing agent.

In a specific embodiment, it may be desirable to administer the pharmaceutical compositions of the invention locally to the area in need of treatment; this may be achieved by, for example, and not by way of limitation, local infusion during surgery, topical application, e.g., in conjunction with a wound dressing after surgery, by injection,

by means of a catheter, by means of a suppository, or by means of an implant, said implant being of a porous, non-porous, or gelatinous material, including membranes, such as sialastic membranes, or fibers. In one embodiment, administration can be by direct injection at the site (or former site) of a malignant tumor or neoplastic or pre-neoplastic tissue.

In another embodiment, the compound can be delivered in a vesicle, in particular a liposome (see Langer, 1990, Science 249:1527-1533; Treat et al., in Liposomes in the Therapy of Infectious Disease and Cancer, Lopez-Berestein and Fidler (eds.), Liss, New York, pp. 353-365 (1989); Lopez-Berestein, ibid., pp. 317-327; see generally ibid.)

In yet another embodiment, the compound can be delivered in a controlled release system. In one embodiment, a pump may be used (see Langer, supra; Sefton, 1987, CRC Crit. Ref. Biomed. Eng. 14:201; Buchwald et al., 1980, Surgery 88:507; Saudek et al., 1989, N. Engl. J. Med. 321:574). In another embodiment, polymeric materials can be used (see Medical Applications of Controlled Release, Langer and Wise (eds.), CRC Pres., Boca Raton, Florida (1974); Controlled Drug Bioavailability, Drug Product Design and Performance, Smolen and Ball (eds.), Wiley, New York (1984); Ranger and Peppas, J., 1983, Macromol. Sci. Rev. Macromol. Chem. 23:61; see also Levy et al., 1985, Science 228:190; During et al., 1989, Ann. Neurol. 25:351; Howard et al., 1989, J. Neurosurg. 71:105). In yet another embodiment, a controlled release system can be placed in proximity of the therapeutic target, i.e., metastatic tissue, thus requiring only a fraction of the systemic dose (see, e.g., Goodson, in Medical Applications of Controlled Release, supra, vol. 2, pp. 115-138 (1984)).

Other controlled release systems are discussed in the review by Langer (1990, Science 249:1527-1533).

In a specific embodiment where the compound of the invention is a nucleic acid encoding a protein, the nucleic acid can be administered in vivo to promote expression of its encoded protein, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, e.g., by use of a retroviral vector (see U.S. Patent No. 4,980,286), or by direct injection, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, or by administering it in linkage to a homeobox-like peptide which is known to enter the nucleus (see e.g., Joliot et al., 1991, Proc. Natl. Acad. Sci. USA 88:1864-1868), etc. Alternatively, a nucleic acid can be introduced

intracellularly and incorporated within host cell DNA for expression, by homologous recombination.

The present invention also provides pharmaceutical compositions. Such compositions comprise a therapeutically effective amount of a compound, and a pharmaceutically acceptable carrier. In a specific embodiment, the term "pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans. The term "carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic is administered. Such pharmaceutical carriers can be sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water is a preferred carrier when the pharmaceutical composition is administered intravenously. Saline solutions and aqueous dextrose and glycerol solutions can also be employed as liquid carriers, particularly for injectable solutions. Suitable pharmaceutical excipients include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. The composition, if desired, can also contain minor amounts of wetting or emulsifying agents, or pH buffering agents. These compositions can take the form of solutions, suspensions, emulsion, tablets, pills, capsules, powders, sustained-release formulations and the like. The composition can be formulated as a suppository, with traditional binders and carriers such as triglycerides. Oral formulation can include standard carriers such as pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, etc. Examples of suitable pharmaceutical carriers are described in "Remington's Pharmaceutical Sciences" by E.W. Martin. Such compositions will contain a therapeutically effective amount of the compound, preferably in purified form, together with a suitable amount of carrier so as to provide the form for proper administration to the patient. The formulation should suit the mode of administration.

In a preferred embodiment, the composition is formulated in accordance with routine procedures as a pharmaceutical composition adapted for intravenous administration to human beings. Typically, compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. Where necessary, the composition may also include a

solubilizing agent and a local anesthetic such as lidocaine to ease pain at the site of the injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the composition is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients may be mixed prior to administration.

The compounds of the invention can be formulated as neutral or salt forms. Pharmaceutically acceptable salts include those formed with free amino groups such as those derived from hydrochloric, phosphoric, acetic, oxalic, tartaric acids, etc., and those formed with free carboxyl groups such as those derived from sodium, potassium, ammonium, calcium, ferric hydroxides, isopropylamine, triethylamine, 2-ethylamino ethanol, histidine, procaine, etc.

The amount of the compound of the invention which will be effective in the treatment of breast cancer can be determined by standard clinical techniques. In addition, in vitro assays may optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness of the disease or disorder, and should be decided according to the judgment of the practitioner and each patient's circumstances. However, suitable dosage ranges for intravenous administration are generally about 20-500 micrograms of active compound per kilogram body weight. Suitable dosage ranges for intranasal administration are generally about 0.01 pg/kg body weight to 1 mg/kg body weight. Effective doses may be extrapolated from dose-response curves derived from in vitro or animal model test systems.

Suppositories generally contain active ingredient in the range of 0.5% to 10% by weight; oral formulations preferably contain 10% to 95% active ingredient.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Optionally associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of

pharmaceuticals or biological products, which notice reflects (a) approval by the agency of manufacture, use or sale for human administration, (b) directions for use, or both.

5. EXAMPLE

5.1 MATERIALS AND METHODS

5.1.1 Preparation of purified cells from reduction mammoplasties

Tissues used in this study consisted of ten samples of reduction mammoplasties carried out for cosmetic reasons from patients aged between 20 and 47 years (mean age ~ 25 yrs). No malignancy or pathology other than minimal fibrocystic change (2 cases) was detected. Purified populations of normal human breast luminal and myoepithelial cells were prepared as described. Clarke et al. (1994) *Epithelial Cell Biol.* 3:38-46. Certain modifications were included to enhance purity. After establishment of primary epithelial cultures over 7-10 days, overnight treatment in calcium-free medium resulted in suspensions enriched in luminal cells. These were incubated for 40 minutes on ice with a mixture of a rat monoclonal antibody (ICR-2, 10μg/ml) against the luminal epithelial marker EMA, and a mouse monoclonal antibody (DAKO, clone SS2/36, 1:25) against the myoepithelial antigen CD-10. The cells were then washed in L-15/10% FCS and labelled for 15 minutes with anti-rat MACS magnetic beads (Miltenyi Biotech Inc), and positively labelled cells separated using a Vario-MACS high field intensity magnet. To obtain a pure population of luminal cells, they were further incubated for 20 minutes with anti-mouse Dynbeads (Dynal, UK) which binds to any residual CD-10 +ve myoepithelial cells, and separated using an MPC-10 low intensity magnet, which does not attract MACS-bead labelled cells. Purified myoepithelial cells were prepared by trypsinisation and filtration (35μM) of the luminally-depleted primary cultures and incubation with a mixture of a mouse IgG2a anti CD-10 antibody (Harlan Serlab, clone 55; 1:50), and a mouse IgG1 monoclonal antibody (F-19, 10μg/ml) reacting with the Fibroblast Activation Protein. Rettig et al. (1993) *Cancer Res.* 53:3327-3335. The labelled suspension of cells was then incubated with sub-class specific anti-mouse IgG2 MACS beads, separated using Vario-MACS, and then incubated with anti-IgG1 mouse Dynabeads to remove any FAP positive cells. This double-labelling procedure enabled F-19 positive fibroblasts, which are the main potential contaminant of the myoepithelial preparation due to their ability to acquire expression of CD-10 antigen in short-term

culture, to be removed from the F-19 negative myoepithelial cells. Using these procedures, purified luminal and myoepithelial cells were obtained in yields of 5×10^6 – 2×10^7 cells from primary epithelial cultures. Purity of the resulting cell preparations was assessed by staining for cell-type specific filament proteins (cytokeratins and vimentin) as originally described by using FACS sorted cells. O'Hare et al (1991) *Differentiation* 46:209-221. The purified cell populations were subsequently washed five times in L-15 medium (serum-free), flash-frozen, and stored at –80°C.

400 µl of the following buffer was then added to each sample:

2M Thiourea (BDH 10348 3D)

8M urea (BDH 45204 3W)

4% CHAPS (Sigma C3023)

65mM dithiotheitol (DTT)

2% (v/v) Resolytes 3.5-10 (BDH 44338 2X)

This mixture was vortexed, centrifuged at 13000 rpm for 5 mins at 15°C, and the supernatant aspirated and stored at –80°C. An assay for protein content was carried out on the final sample (Pierce BCA Cat # 23225).

5.1.2 Isoelectric Focusing

Isoelectric focusing (IEF), was performed using the Immobiline® DryStrip Kit (Pharmacia BioTech), following the procedure described in the manufacturer's instructions, see Instructions for Immobiline® DryStrip Kit, Pharmacia, # 18-1038-63, Edition AB (incorporated herein by reference in its entirety). Immobilized pH Gradient (IPG) strips (18cm, pH 3-10 non-linear strips; Pharmacia Cat. # 17-1235-01) were rehydrated overnight at 20°C in a solution of 8M urea, 2% (w/v) CHAPS, 10mM DTT, 2% (v/v) Resolytes 3.5-10, as described in the Immobiline DryStrip Users Manual. For IEF, 50µl of supernatant (prepared as above) was loaded onto a strip, with the cup-loading units being placed at the basic end of the strip. The loaded gels were then covered with mineral oil (Pharmacia 17-3335-01) and a voltage was immediately applied to the strips according to the following profile, using a Pharmacia EPS3500XL power supply (Cat 19-3500-01):

Initial voltage = 300V for 2 hrs

Linear Ramp from 300V to 3500V over 3hrs

Hold at 3500V for 19hrs

For all stages of the process, the current limit was set to 10mA for 12 gels, and the wattage limit to 5W. The temperature was held at 20°C throughout the run.

5.1.3 Gel Equilibration and SDS-PAGE

After the final 19hr step, the strips were immediately removed and immersed for 10 mins at 20°C in a first solution of the following composition: 6M urea; 2% (w/v) DTT; 2% (w/v) SDS; 30% (v/v) glycerol (Fluka 49767); 0.05M Tris/HCl, pH 6.8 (Sigma Cat T-1503). The strips were removed from the first solution and immersed for 10 mins at 20°C in a second solution of the following composition: 6M urea; 2% (w/v) iodoacetamide (Sigma I-6125); 2% (w/v) SDS; 30% (v/v) glycerol; 0.05M Tris/HCl, pH 6.8. After removal from the second solution, the strips were loaded onto supported gels for SDS-PAGE according to Hochstrasser et al., 1988, Analytical Biochemistry 173: 412-423 (incorporated herein by reference in its entirety), with modifications as specified below.

5.1.4 Preparation of supported gels

The gels were cast between two glass plates of the following dimensions: 23cm wide x 24cm long (back plate); 23cm wide x 24cm long with a 2cm deep notch in the central 19cm (front plate). To promote covalent attachment of SDS-PAGE gels, the back plate was treated with a 0.4% solution of alpha-methacryl-oxypropyltrimethoxysilane in ethanol (BindSilan; Pharmacia Cat. # 17-1330-01). The front plate was treated with RepelSilane (Pharmacia Cat. # 17-1332-01) to reduce adhesion of the gel. Excess reagent was removed by washing with water, and the plates were allowed to dry. At this stage, both as identification for the gel, and as a marker to identify the coated face of the plate, an adhesive bar-code was attached to the back plate in a position such that it would not come into contact with the gel matrix.

The dried plates were assembled into a casting box with a capacity of 13 gel sandwiches. The top and bottom plates of each sandwich were spaced by means of 1mm thick spacers, 2.5 cm wide. The sandwiches were interleaved with acetate sheets to facilitate separation of the sandwiches after gel polymerization. Casting was then carried out according to Hochstrasser et al., op. cit.

A 9-16% linear polyacrylamide gradient was cast, extending up to a point 2cm below the level of the notch in the front plate, using the Angelique gradient casting system (Large Scale Biology). Stock solutions were as follows. Acrylamide (40% in water) was from Serva (Cat. # 10677). The cross-linking agent was PDA (BioRad 161-0202), at a concentration of 2.6% (w/w) of the total starting monomer content. The gel buffer was 0.375M Tris/HCl, pH 8.8. The polymerization catalyst was 0.05% (v/v) TEMED (BioRad 161-0801), and the initiator was 0.1% (w/v) APS (BioRad 161-0700). No SDS was included in the gel and no stacking gel was used. The cast gels were allowed to polymerize at 20°C overnight, and then stored at 4°C in sealed polyethylene bags with 6ml of gel buffer, and were used within 4 weeks.

5.1.5 SDS-PAGE

The gels were placed in a running tank, as described by Amess et al., 1995, Electrophoresis 16: 1255-1267 (incorporated herein by reference in its entirety). The tank was filled with running buffer (as above) until the level of the buffer was just higher than the top of the region of the gels which contained polyacrylamide, so as to achieve efficient cooling of the active gel area. Running buffer was added to the top buffer compartments formed by the gels, and then voltage was applied immediately to the gels using a Consort E-833 power supply. For 1 hour, the gels were run at 20mA/gel. The wattage limit was set to 150W for a tank containing 6 gels, and the voltage limit was set to 600V. After 1 hour, the gels were then run at 40mA/gel, with the same voltage and wattage limits as before, until the bromophenol blue line was 0.5cm from the bottom of the gel. The temperature of the buffer was held at 10°C throughout the run.

5.1.6 Staining

Upon completion of the electrophoresis run, the gels were immediately removed from the tank for fixation. The top plate of the gel cassette was carefully removed, leaving the gel bonded to the bottom plate. The bottom plate with its attached gel was then placed into a staining apparatus, which can accommodate 12 gels. The gels were completely immersed in fixative solution of 40% (v/v) ethanol (BDH 28719), 10% (v/v) acetic acid (BDH 100016X), 50% (v/v) water (MilliQ-Millipore), which was continuously circulated over the gels. After an overnight incubation, the fixative was drained from the tank, and the gels were primed by immersion in 7.5% (v/v) acetic acid, 0.05% (w/v) SDS, 92.5% (v/v)

water for 30 mins. The priming solution was then drained, and the gels were stained by complete immersion in a staining solution for 4 hours (e.g. Sypro Red, Molecular Probes Inc., Eugene, Oregon). A preferred fluorescent dye is described in U.S. Application No. 09/412,168, filed on October 5, 1999, which is incorporated herein by reference in its entirety.

5.1.7 Imaging of the gel

A computer-readable output was produced by imaging the fluorescently stained gels with an Apollo 2 scanner (Oxford Glycosciences, Oxford, UK, described above). The gels were removed from the stain, rinsed with water briefly, and imaged on the Apollo 2 Scanner, in Red Fluorescence mode with a PMT setting of 1000V, and a resolution of 200μm (microns). Since the gel was rigidly bonded to a glass plate, the gel was held in contact with the scanner bed during imaging. To avoid interference patterns arising from non-uniform contact between the gel and the scanner bed, a film of water was introduced under the gel, taking care to avoid air pockets. Moreover, the gel was placed in a frame provided with two fluorescent buttons that were imaged together with the gel to provide reference points (designated M1 and M2) for determining the x,y coordinates of other features detected in the gel. A matched frame was provided on a robotic gel excisor in order to preserve accurate alignment of the gel. After imaging, the gels were sealed in polyethylene bags containing a small volume of staining solution, and then stored at 4°C.

5.1.8 Digital Analysis of the Data

The data were processed as described in U.S. Application Serial No. 08/980,574, (published as WO 98/23950) Sections 5.4 and 5.5 (incorporated herein by reference), as set forth more particularly below.

The output from the scanner was first processed using the MELANIE® II 2D PAGE analysis program (Release 2.2, 1997, BioRad Laboratories, Hercules, California, Cat. # 170-7566) to autodetect the registration points, M1 and M2; to autocrop the images (i.e., to eliminate signals originating from areas of the scanned image lying outside the boundaries of the gel, e.g. the reference frame); to filter out artifacts due to dust; to detect and quantify features; and to create image files in GIF format. Features were detected using the following parameters:

Smooths =2

Laplacian threshold 50
Partials threshold 1
Saturation = 100
Peakedness = 0
Minimum Perimeter = 10

5.1.9 Assignment of pI and MW Values

All images for the gels in the data set were calibrated with respect to pI and MW using a standard breast tissue image (from a pooled sample) which had itself been calibrated with respect to a standard E. coli calibration sample. Optimally 12 or at least 10 or more such calibration landmarks were selected on the basis that (a) they were protein features expected to be found in each breast tissue sample, (b) they were broadly located across a wide range of pI and MW, and (c) each was a known protein with known pI and MW. The calibration landmarks were reported in each of the study gels by comparison with the calibration gel. The Melanie-II software described above then used these landmarks to interpolate the pI and MW values for each of the other protein features n the study gels. Images were edited to remove gross artifacts such as dust, to reject images which had gross abnormalities such as smearing of protein features, or were of too low a loading or overall image intensity to allow identification of more than the most intense features. Images were then compared by pairing with one common image from the whole sample set. This common image, a primary master image, was selected on the basis of protein load (maximum load consistent with maximum feature detection), a well resolved myoglobin region, (myoglobin was used as an internal standard) general quality and how representative the image was of all those to be included in the analysis. The set of all primary master images was analysed for background CV. Images exceeding this CV were rejected.

Each of the remaining images was individually matched to the primary master image such that common protein features were paired between them.

Using the paired protein features as a guide, the geometry of each of the gels was adjusted to obtain maximum alignment between its pattern of protein features, and that of the primary master. All the images were added together to create a composite master image, and the positions and shapes of all the gel features of all the component images were super-imposed into this composite master.

Overlapping sets of features in the composite master were grouped together to form individual master features. Each master feature, representing a set of matched features from one or more of the component gel images, was given a unique identifying index, the molecular cluster index (MCI). Certain master features on the periphery of the composite master image that were outside the well resolved area of the gel were excluded from further analysis.

5.1.10 Construction of Profiles

After matching all gels in the dataset to the final master group, the intensity of each feature was measured and stored. The end result of this aspect of the analysis was the generation of a digital profile which contained, for each identified spot: 1) a unique arbitrary identification code, 2) the x,y coordinates, 3) the isoelectric point, 4) the molecular weight, 5) the signal value, 6) the standard deviation for each of the preceding measurements, and 7) a pointer to the MCI of the spot on the master gel to which this spot was matched. By virtue of a Laboratory Information Management System (LIMS), this profile was traceable to the actual stored gel from which it was generated, so that proteins identified by computer analysis of gel profile databases could be retrieved. The LIMS also permitted the profile to be traced back to the original sample or patient.

5.1.11 Cross-matching Between Samples

Once the profile was generated, analysis was directed toward the selection of interesting proteins. Each significant feature in a profile was assigned an index, the "Molecular Cluster Index" (MCI) that identifies the molecular content of the feature and has the same value in matching features in all gels. A molecular cluster table was generated from the master gel for each sample type (i.e. luminal epithelial cell breast tissue and myoepithelial cell breast tissue). Gels from all other samples of the same type were matched with the relevant master gel. The digital profiles for each sample were then annotated by adding, for each matched feature, the MCI assigned to that feature in the master profile.

5.1.12 Differential Analysis of the Profiles

Within each sample set (luminal epithelial cell breast tissue and myoepithelial cell breast tissue), the profiles were analyzed to identify and select those features present in at least

80% of the profiles. These selected features were then assembled into a luminal epithelial cell feature set and a myoepithelial cell feature set. Differentially present features were identified as Breast Cancer-Associated Features (BFs).

5.1.13 Recovery and analysis of selected proteins

Proteins in BFs were robotically excised as described in U.S. Application Serial No. 08/980,574, (published as WO 98/23950) Sections 5.6 and 5.7, and processed to generate tryptic peptides; partial amino acid sequences of these peptides were determined by mass spectroscopy, using de novo sequencing, as described in Application No. 08/877,605, filed June 18, 1997 (published as WO98/53323) and Application No. 09/094,996, filed June 15, 1998, each of which is incorporated herein by reference in its entirety.

5.2 RESULTS

These initial experiments identified 33 features that were decreased and 18 features that were increased in sets of matched normal breast luminal cells, purified from normal breast tissue of living human subjects, as compared with corresponding sets of matched purified normal human myoepithelial cells from the same subjects. Details of these BFs are provided in Tables I and II.

Partial amino acid sequences were determined for the differentially present BPIs in these BFs. Details of these BPIs are provided in Tables III and IV.

CLAIMS

1. A method for screening for and/or diagnosis of breast cancer in a human subject, which comprises the step of identifying the presence or absence of one or more of the protein features as defined in tables I, II, III and IV herein, in a biological sample obtained from said human subject.

2. A method for monitoring and/or assessing breast cancer treatment in a human subject, which comprises the step of identifying the presence or absence of one or more of the protein features as defined in tables I, II, III and IV herein, in a biological sample obtained from said human subject.

3. A method for identifying the presence or absence of metastatic breast cancer cells in a biological sample obtained from a human subject, which comprises the step of identifying the presence or absence of one or more of the protein features as defined in tables I, II, III and IV herein.

4. A method as claimed in any one of claims 1 to 3 wherein the biological sample is a serum sample or a tissue sample.

5. A method as claimed in any one of claims 1 to 4 wherein a "cluster" or subset of the totality of protein features defined in tables I, II, III and IV is identified.

6. A method as claimed in any one of claims 1 to 5 wherein the method comprises an immunoassay step utilising one or more antibodies against one or more of the protein features defined in tables I, II, III and IV, or a derivative, homologue or fragment thereof.

7. A method as claimed in claim 6 wherein the immunoassay is a competitive immunoassay, a non-competitive assay system using techniques such as western blots, a radioimmunoassay, an ELISA (enzyme linked immunosorbent assay), a "sandwich"

immunoassay, an immunoprecipitation assay, a precipitin reaction, a gel diffusion precipitin reaction, an immunodiffusion assay, an agglutination assay, a complement-fixation assay, an immunoradiometric assay, a fluorescent immunoassay , a protein A immunoassay, ab Immunoprecipitation assay or an Immunohistochemical assay.

8. A method as claimed in any one of claims 1 to 5 wherein the method comprises the use of nucleic acid probes and/or PCR reactions to amplify nucleic acid coding for one or more of the protein features defined in tables I, II, III and IV.

9. An antibody which specifically binds to a protein feature as defined in tables I, II, III or IV.

10. An antibody as claimed in claim 9 which is for use in the screening for and/or diagnosis of breast cancer in a human subject.

11. An antibody as claimed in claim 9 or claim 10 which is a monoclonal antibody.

12. An antibody as claimed in any one of claims 9 to 11 wherein the antibody is adapted/modified such that binding to the protein will be localised to the site of the breast cancer cells, preferably metastatic breast cancer cells.

13. A diagnostic kit comprising one or more reagents for use in the detection and/or determination of one or more of the protein features defined in tables I, II, III and IV.

14. A kit as claimed in claim 13 which comprises one or more containers with one or more antibodies against one or more of the protein features.

15. A kit as claimed in claim 14 which further comprises a labeled binding partner to the antibody and/or a solid phase (such as a reagent strip) upon which the antibody(ies) is/are immobilized.

16. A kit as claimed in claim 13 which comprises a nucleic acid probe capable of hybridizing to DNA or RNA encoding one or more of the protein features defined in Tables I, II, III or IV.

17. A kit as claimed in claim 16 which comprises in one or more containers a pair of primers (*e.g.*, each in the size range of 6-30 nucleotides, more preferably 10-20 nucleotides) that are capable of priming amplification under appropriate reaction conditions of at least a portion of a nucleic acid encoding a protein feature as defined in Tables I, II, III or IV.

18. A method as claimed in any one of claims 1 to 5 wherein a whole body or organ scan of the subject is carried out to determine localisation of breast tissue cells, particularly metastatic breast cancer cells.

19. A method as claimed in claim 18 wherein labelled antibodies are employed.

20. A method as claimed in claim 19 wherein the antibodies are radiolabelled.

21. A method for the treatment of breast cancer, particularly metastatic breast cancer, which comprises administering to a subject suffering from said breast cancer, an effective amount of one or more antibodies against one or more of the protein features defined in Tables I, II, III or IV, in association with or conjugated to an agent capable of causing cell death.

22. A method as claimed in claim 21 wherein the agent is a cytotoxic agent or a cytostatic agent.

23. The use of one or more antibodies against one or more of the protein features defined in Tables I, II, III or IV, in association with or conjugated to an agent capable of

causing cell death in the manufacture of a medicament for the treatment of breast cancer, particularly metastatic breast cancer.

INTERNATIONAL SEARCH REPORT

Inte onal Application No
PCT/GB 00/00908

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G01N33/574 C07K16/30 C12Q1/68 G01N33/561

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G01N C07K C12Q A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 188 964 A (MCGUIRE WILLIAM L ET AL) 23 February 1993 (1993-02-23) tables 3A,3B ---	1-7, 9-15, 18-20
X	US 5 798 266 A (QUAY STEVEN C ET AL) 25 August 1998 (1998-08-25) example 9 ---	1-7, 9-15, 18-20
X	US 4 775 620 A (CARDIFF ROBERT D ET AL) 4 October 1988 (1988-10-04) abstract --- -/--	1-7, 9-15, 18-20

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
7 July 2000	24/07/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hart-Davis, J

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inte ional Application No
PCT/GB 00/00908

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 158 893 A (HACKETT ADELINE J ET AL) 27 October 1992 (1992-10-27) the whole document	1-7, 9-15, 18-20
X	WO 96 17080 A (IMP CANCER RES TECH ;SELBY PETER JOHN (GB); BURCHILL SUSAN ANN (GB) 6 June 1996 (1996-06-06) the whole document	1,8,13, 16,17
X	BINI LUCA; MAGI BARBARA ET AL.: "Protein expression profiles in human breast ductal carcinoma and histologically normal tissue" ELECTROPHORESIS, vol. 18, December 1997 (1997-12), pages 2832-2841, XP000923292 the whole document	1-5
X	WILLIAMS KATHERINE; CHUBB CYNTHIA; HUBERMAN ELIEZER; GIOMETTI CAROL S: "Analysis of differential protein expression in normal and neoplastic human breast epithelial cell lines" ELECTROPHORESIS, vol. 19, February 1998 (1998-02), pages 333-343, XP000923169 the whole document	1-5
X	FRANZEN BO; LINDER STIG ET AL.: "Analysis of polypeptide expression in benign and malignant human breast lesions" ELECTROPHORESIS, vol. 18, 1997, pages 582-587, XP000923187 the whole document	1-5
A	WO 97 41441 A (FREEMAN MICHAEL R ;WIEDERSCHAIN DMITRI (US); MOSES MARSHA A (US);) 6 November 1997 (1997-11-06) claims 1,50-52; tables 1-7	1-5
P,X	BERGMAN ANN-CHARLOTTE; BENJAMIN TIMOTHY; ET AL.: "Identification of gel-separated tumor marker proteins by mass spectrometry" ELECTROPHORESIS, vol. 21, February 2000 (2000-02), pages 679-686, XP000923086 the whole document	1-5

-/--

1

INTERNATIONAL SEARCH REPORT

Inte .onal Application No
PCT/GB 00/00908

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
P,X	PAGE MARTIN J; AMESS BOB ET AL.: "Proteomic definition of normal human luminal and myoepithelial breast cells purified from reduction mammoplasties" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, vol. 96, 26 October 1999 (1999-10-26), pages 12589-12594, XP002142084 the whole document	1-5

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. ional Application No
PCT/GB 00/00908

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5188964 A	23-02-1993	AU 7785091 A	11-11-1991
		CA 2079345 A	13-10-1991
		EP 0525115 A	03-02-1993
		WO 9116632 A	31-10-1991
		US 5447843 A	05-09-1995
US 5798266 A	25-08-1998	AU 4085097 A	19-03-1998
		EP 0932699 A	04-08-1999
		WO 9808976 A	05-03-1998
US 4775620 A	04-10-1988	AT 125626 T	15-08-1995
		CA 1336171 A	04-07-1995
		DE 3588043 D	31-08-1995
		DE 3588043 T	21-03-1996
		EP 0167616 A	15-01-1986
		JP 2617289 B	04-06-1997
		JP 8275778 A	22-10-1996
		JP 1047392 A	21-02-1989
		JP 2537539 B	25-09-1996
		JP 1052800 A	28-02-1989
		JP 1955428 C	28-07-1995
		JP 6081760 B	19-10-1994
		JP 2896362 B	31-05-1999
		JP 10123138 A	15-05-1998
		JP 2723203 B	09-03-1998
		JP 61501048 T	22-05-1986
		WO 8503132 A	18-07-1985
US 5158893 A	27-10-1992	NONE	
WO 9617080 A	06-06-1996	NONE	
WO 9741441 A	06-11-1997	AU 2681997 A	19-11-1997
		CA 2252648 A	06-11-1997
		EP 0900381 A	10-03-1999
		US 6037138 A	14-03-2000